Uso Não Financeiro de *Blockchain*: Um Estudo de Caso Sobre o Registro, Autenticação e Preservação de Documentos Digitais Acadêmicos

Rostand Costa[1], **Daniel Faustino**[1], **Guido Lemos**[1], **Ademir Queiroga**[1], **Cláudio Djohnnatha**[1], **Felipe Alves**[1], **Jordan Lira**[1] **e Mateus Pires**[1]

[1]Laboratório de Aplicações de Vídeo Digital - LAVID
Centro de Informática - UFPB
Campus V - Mangabeira – João Pessoa – PB – Brazil – Caixa Postal 58051-900

Abstract. *The purpose of this paper is to to investigate the potential use of the blockchain technology combined with active distributed repositories to create a platform, scalable and agnostic*[1]*, specialized in the authentication and preservation of digital documents. As a proof of concept of the proposed platform, was performed the construction of a public service for digital registration and verification of the authenticity of academic documents. The prototype of the service offers an interface for educational institutions to register official documents, such as diplomas and certificates, using blockchain and an interface so that users can verify the authenticity of a document through its registration number in a DLT. The documents registered in the service are automatically inserted in the long-term digital preservation repository.*

1. Introdução

A validação da existência ou da posse de documentos formalmente assinados é fundamental em qualquer contexto legal. Normalmente, a certificação tradicional de documentos físicos se baseia em autoridades centrais, notariais ou não, para armazenar e aplicar os registros e mecanismos necessários para tal fim e também lidar com os aspectos e desafios da segurança. Desafios esses que se tornam cada vez mais difíceis à medida que os arquivos envelhecem.

Entretanto, a materialização e desmaterialização de documentos bem como o dinamismo e velocidade das relações digitais têm representado uma nova frente para entidades produtoras e/ou certificadoras de documentos. Principalmente quando começa a emergir a possibilidade de geração de documentos em papel a partir de documentos digitais e a geração de documentos digitais a partir de documentos em papel, demandando a garantia de que os termos estabelecidos no original sejam efetivamente conservados e recebam uma chancela de legitimação, independentemente da sua forma de representação.

De um ponto de vista prático, a certificação de documentos digitais apresenta três dimensões principais [Crosby et al. 2016]: i) *Prova de Propriedade/Autoria* (quem é o detentor/autor do documento), ii) *Prova de Integridade* (o documento está íntegro e exatamente igual à quando foi criado) e iii) *Prova de Existência* (o documento foi criado e legitimado em um dado momento no tempo).

Neste sentido, a tecnologia de livros-razão distribuídos [Kakavand et al. 2017] (ou DLTs, do inglês *Distributed Ledgers Technologies*), normalmente baseados em *blockchain*,

[1]Agnóstica, no contexto deste documento, significa independência do formato de representação e do tipo de conteúdo armazenado no objeto digital a ser registrado e preservado.

se apresenta como um modelo alternativo para a certificação de documentos legais, sobretudo pela eliminação da necessidade de uma autoridade centralizada para verificar a autenticidade de um documento. Uma entidade emissora pode simplesmente armazenar a assinatura e a marcação de tempo associada com um documento legal na cadeia de blocos e validá-lo em qualquer tempo usando os mecanismos nativos da tecnologia [Wikipedia 2018].

Este artigo traz os resultados preliminares de um projeto de pesquisa e desenvolvimento[2] cujo objetivo é investigar o potencial do uso combinado das tecnologias de *blockchain*, certificação digital e preservação digital para a criação de uma plataforma, escalável e agnóstica, especializada na autenticação e preservação de documentos digitais.

Como prova de conceito da plataforma proposta, foi feita a construção de um serviço público para registro e verificação digital da autenticidade de documentos acadêmicos. O protótipo de serviço oferece uma interface para que instituições de ensino possam registrar documentos oficiais, como diplomas e certificados, usando *blockchain* e uma interface para que outras instituições e/ou interessados possam verificar a legitimidade de um documento através do seu número de registro em uma DLT. Os documentos registrados no serviço são automaticamente inseridos em um repositório de preservação digital de longo termo.

O restante do documento está organizado como segue. A Seção 2 traz a contextualização e motivação do projeto, incluindo uma breve discussão sobre ocorrências de fraudes associadas com diplomas acadêmicos, como a gestão de tais documentos ocorre e como anda o processo de adoção de diplomas digitais. Na Seção 3 é feita a apresentação da plataforma proposta e também é discutida a estratégia de modelagem adotada, de forma a permitir conciliar a evolução tecnológica sugerida com as transformações culturais e regulatórias necessárias. Na Seção 4 é realizado o detalhamento de como o protótipo da plataforma foi projetado e construído e alguns resultados preliminares de integração com duas IES de grande porte, uma pública (UFPB) e uma privada (PUC-RJ). A Seção 5, finalmente, contém nossas considerações finais.

2. Contextualização e Motivação

2.1. Necessidade de Proteção de Documentos Acadêmicos

Os últimos dados[3] do Censo da Educação Superior realizado anualmente pelo INEP indicam que o Brasil teve em 2016 um total de um milhão e cem mil concluintes e um total de quase três milhões de ingressantes nas instituições públicas e privadas de ensino superior. Com isso, em 2016, o total de estudantes matriculados em cursos de ensino superior no país ultrapassa pela primeira vez os 8 milhões de alunos.

Os números[4] da CAPES sobre os programas de pós-graduação no Brasil indicam que em 2016 haviam 325.320 estudantes matriculados em cursos de pós-graduação *stricto sensu* no país, variando entre mestrado profissionalizante, mestrado acadêmico e doutorado. Neste mesmo ano foram titulados no Brasil 20.630 doutores e 59.349 mestres.

Em relação aos cursos de pós-graduação *lato sensu*, não há dados atualizados sobre a quantidade de cursos em funcionamento ou sobre a quantidade de alunos matriculados e titulados. São números bastante variáveis dada a natureza mais dinâmica desses cursos, sempre

[2] $http://gt-rap.lavid.ufpb.br$
[3] $http://portal.inep.gov.br$
[4] $http://www.capes.gov.br/$

associados a necessidades do mercado, e também devido as baixas exigências regulamentares para sua abertura. Porém, é possível estimar, de forma bastante conservadora, que a pós-graduação *lato sensu* no Brasil seja, pelo menos, 10 vezes maior que a pós-graduação *stricto sensu*, o que nos levaria a diplomação de pelo menos 600.000 especialistas ao ano.

Com isso, podemos chegar a conclusão que no Brasil, apenas no ensino superior, são emitidos anualmente cerca de 1,8 milhão de diplomas. Se somarmos a este número a quantidade de diplomas emitidos no exterior e revalidados no país, é possível atingirmos a marca de 2 milhões de diplomas emitidos anualmente.

Apesar da emissão de tais diplomas ser controlada, exigindo-se o registro do documento por uma Universidade junto ao Ministério da Educação, as dimensões continentais de nosso país, aliadas à falta de suporte tecnológico, fazem com que a tarefa de verificar a autenticidade de um diploma emitido ou revalidado no país seja feita de forma ineficiente.

Um exemplo dessa dificuldade envolveu o INEP, um órgão do próprio Ministério da Educação, responsável por diversas atividades relacionadas a avaliação do ensino no pais. Em 2011 o INEP foi obrigado a cancelar as avaliações de 4 cursos de graduação na área de Direito pelo fato de que um dos avaliadores *ad hoc* participantes do **Banco de Avaliadores do Sistema Nacional de Educação Superior** (BASIS) teria apresentado diplomas falsos de mestrado e doutorado [FOLHA 2011]. Tal fraude só foi descoberta por conta de uma denúncia anônima. Não fosse isso, possivelmente este avaliador teria continuado a representar o Ministério da Educação na avaliação do ensino superior no país.

Casos de uso de diplomas falsos para ingresso no ensino superior e no serviço público também não são raros. O Ministério Público do Paraná investigou o uso de mais de 500 diplomas falsos na cidade de Maringá que foram utilizados para ingressos em universidades e aprovação em concursos públicos [do Povo 2013].

Há quadrilhas especializadas em vender diplomas falsos em nome de Instituições de Ensino Superior [GLOBO 2017]. No estado do Espírito Santo, mais de 100 professores estão sendo processados por usarem diplomas falsos [Online 2017]. As fraudes também acontecem no processo de revalidação de diplomas estrangeiros. Um caso emblemático, foi a quadrilha que revalidava de forma fraudulenta diplomas de medicina no estado do Mato Grosso [GLOBO 2013]. Além disso, há inúmeros relatos de casos de utilização de diplomas falsos para ludibriar clientes e, o que é mais grave, pacientes, no caso dos diplomas relacionados às áreas de saúde.

O processo de investigação, descoberta e tomada de providência em relação à falsificação de diplomas também envolve desafios específicos. Como os mecanismos de controles e verificação nem sempre são automatizados e o número de casos têm crescido nos últimos anos, é complicado para a justiça detectar e tomar as devidas providências frente aos prejuízos causados por essas quadrilhas e pessoas de má fé. A Secretaria de Estado da Educação do Espírito Santo (SEDU-ES), por exemplo, relata que as investigações referentes aos processos administrativos para apurar o uso de diplomas falsos nas escolas do estado podem durar até 180 dias, com possibilidade de prorrogação [Online 2017].

2.1.1. Gestão de Diplomas Acadêmicos

A *LDB - Lei de Diretrizes e Bases da Educação* (Lei 9394/1996) delega a responsabilidade pela emissão dos diplomas para as Instituições de Ensino Superior, além dos demais do-

cumentos acessórios como declaração de conclusão e certificados de conclusão de curso. Além disso, segundo a própria LDB, a responsabilidade pelo registro e manutenção de tais registros também é de responsabilidade da IES. Para o caso de instituições de ensino não-universitárias, tais registros são feitos por instituições indicadas pelo Conselho Nacional de Educação (CNE).

Quanto a revalidação de diplomas de universidades do exterior e emissão e registro de diplomas de pós-graduação, tais responsabilidades também são das instituições universitárias brasileiras e seguem as mesmas regras dos diplomas de graduação, observando a competência da universidade em relação àquela área.

Em relação à preservação dos registros, as universidades devem manter os registros e documentos emitidos por elas, relacionados à instituição em si e às faculdades e demais instituições de ensino que não possuem competência para registro.

A gestão de tais documentos está regulamentada pela Lei 8159/91. A referida lei cria o *Conselho Nacional de Arquivos* (CONARQ). O CONARQ estabelece recomendações e define a tabela de temporalidade de manutenção de arquivos públicos pelos órgãos responsáveis. Os diplomas universitários (código 135.421) possuem tempo de guarda recomendada de 5 anos com posterior eliminação. Já o registro do diploma possui tempo de guarda ativo de 5 anos, com posterior guarda permanente. No caso das instituições particulares o registro é efetuado por instituição credenciada e a documentação base é retornada para instituição não credenciada a fim de que proceda com a preservação.

Ainda sobre os diplomas de faculdades ofertantes não credenciadas, segundo a norma técnica 391/2013/MEC, cabe às universidades credenciadas apenas registrar o diploma, porém a emissão é de responsabilidade da instituição não credenciada. Neste sentido, cada instituição define seu fluxograma de emissão e registro de diplomas. Geralmente esse fluxo está descrito no regimento interno e nos documentos que definem os processos internos da instituição.

2.1.2. Uso de Diplomas Digitais

O uso de diplomas digitais apresenta uma série de vantagens quando comparado ao diploma tradicional: i) a eliminação, pelo emissor, do custo de impressão da versão em papel (normalmente especial e de alto custo); ii) a replicação e distribuição ilimitada e gratuita do documento pelo portador; e iii) a possibilidade de adoção, pelo destinatário, de mecanismos automatizados de verificação da autenticidade do documento. Tais características emprestam mais agilidade e segurança para que portadores de diplomas e interessados possam compartilhar e verificar a legitimidade de tais documentos de forma mais eficiente. Além disso, é possível padronizar e formalizar o processo de emissão e validação dos diplomas, dificultando tentativas de fraudes.

Embora o uso de diplomas digitais desponte como um caminho quase que natural para a gestão de documentos dessa natureza, ainda não existe uma legislação que trate diretamente do tema. A falta de uma regulação específica, assim como aspectos culturais, atrasam o avanço do processo de digitalização desses documentos e o seu uso em larga escala.

Um análise mais direta sobre isso está contido no parecer CNE/CES No. 226/2012, o qual faz algumas considerações sobre o uso de diplomas digitais a partir de uma consulta feita pela UNIVAP. Resumidamente, o parecer em questão indica que o uso desse tipo de

documento não é proibido, porém a IES deve oferecer também a possibilidade de emissão de cópia física do documento, deixando a critério do aluno a escolha da melhor forma de acesso ao seu diploma.

Em uma iniciativa mais concreta, em 2013 a USP passou a oferecer a emissão de diplomas no formato digital [ITI 2013]. Um fato curioso ocorreu em 2016, quando a USP, mesmo após a adoção do diploma digital, ainda deixou de emitir mais de 4 mil diplomas por falta de papel [Estadao 2016], o que sugere que os dois tipos de diploma ainda conviviam na época. O Centro Universitário de Belas Artes foi outra instituição paulista de ensino que também adotou a emissão de diplomas digitais a partir de 2013 [Baguete 2013].

2.2. Desafios da Preservação Digital de Longo Termo

Na mesma proporção em que uma parte considerável dos artefatos relacionados à diversas atividades humanas está sendo criada em formatos digitais, é esperado que as práticas de preservação para essas informações também devam ser baseadas em técnicas e tecnologias adequadas e igualmente digitais.

Quando comparada com a preservação de coleções físicas, a preservação de conteúdo digital traz, em si, uma associação, quase paradoxal, de um grande potencial de risco e um grande potencial de proteção [Skinner and Schultz 2010]. O potencial de risco é representado pela efemeridade do armazenamento digital que pode ser irremediavelmente perdido por causa de falhas técnicas ou humanas com muito mais facilidade e rapidez do que no caso de representações físicas de conteúdo. O potencial de proteção, por sua vez, é ancorado no fato de que coleções digitais podem ser indefinidamente reproduzidas e armazenadas com total fidelidade e integridade.

A área da preservação digital ainda está nos estágios iniciais de sua formação e o aparato tecnológico, metodológico e político para preservar a informação digital ainda está sendo construído. Boa parte do conhecimento acumulado na última década em preservação e acesso a recursos digitais está se consolidando em um conjunto de estratégias, abordagens tecnológicas e atividades que agora são coletivamente conhecidas como "curadoria digital". Ainda um conceito em evolução, a curadoria digital envolve a gestão atuante e a preservação de recursos digitais durante todo o seu ciclo de interesse, tendo como perspectiva o desafio de longo prazo de atender a gerações atuais e futuras de usuários [Sayão and Sales 2012].

A perfeita continuidade de coleções digitais depende, em grande parte, de se buscar um equilíbrio da aplicação de medidas que aproveitem ao máximo o potencial de proteção ao ponto de neutralizar o seu inerente potencial de risco. Entretanto, o desafio pode representar muito mais um problema social e institucional do que uma questão meramente técnica, pois, em particular, para a preservação digital no meio acadêmico, depende-se de instituições que passam por mudanças de direção, missão, administração e fontes de financiamento[Sayão and Sales 2012].

Além disso, a preservação digital envolve desafios essencialmente diferentes dos encontrados na conservação de conteúdo em suportes mais tradicionais. De um ponto de vista mais tradicional, o ato de preservar traduz-se no ato de manter imutável e intacto. No ambiente digital, entretanto, a ação de preservar também pode se referir a mudar, recriar e renovar. Onde renovar pode significar mudar formatos, atualizar mídias e/ou substituir *hardware* e *software*. Em suma: se, por um lado, queremos manter o conteúdo exatamente como foi criado, intacto; por outro lado, queremos continuar acessando-o em plataformas modernas. Este é o *Paradoxo da Preservação Digital*, conforme descrito por Sayão

[Sayão and Sales 2012].

Um número crescente de organizações de memória cultural (incluindo as reunidas na iniciativa *MetaArchive*[Skinner and Halbert 2009]) aposta que os esforços mais eficazes de preservação digital ocorrem na prática através de alguma estratégia para manter múltiplas cópias de conteúdo digital em locais distribuídos seguros[Ruusalepp and Dobreva 2012, Ferreira et al. 2012]. Na era digital, esta estratégia requer investimentos numa matriz distribuída de servidores, capazes de armazenar coleções digitais em uma metodologia pré-coordenada.

A montagem de infraestruturas computacionais para preservação digital distribuída[Skinner and Schultz 2010] implica em adotar estratégias envolvendo a distribuição geográfica do armazenamento em vários locais e a implementação de segurança forte em caches individuais, uma combinação de abordagens que maximiza a sobrevivência de conteúdo, tanto em termos individuais quanto coletivos. Maximizar as medidas de segurança implementadas em caches individuais reduz a probabilidade de que qualquer cache individual seja comprometida. Por sua vez, a replicação reduz a probabilidade de que a perda de qualquer cache individual leve a uma perda do conteúdo preservado.

Entretanto, é pouco provável que uma única organização educacional tenha a capacidade de operar de forma adequada vários servidores distribuídos geograficamente. Neste sentido, a colaboração entre instituições é essencial, e tal colaboração exige investimentos técnicos e organizacionais [Costa et al. 2015, Costa et al. 2016]. Não é o caso de contar apenas com uma solução tecnológica adequada, mas também precisam ser estabelecidos acordos interinstitucionais robustos de longo prazo, ou não haverá compromisso suficiente para uma atuação sintonizada ao longo do tempo.

3. Serviço de Autenticação e Preservação Digital de Documentos Acadêmicos

É baseado neste cenário que propomos a construção de uma plataforma, escalável e agnóstica, para armazenamento e verificação digital da autenticidade de documentos digitais baseada no uso combinado das tecnologias de *blockchain*, certificação digital e preservação digital.

A estratégia utilizada como metodologia para o desenvolvimento do projeto previu a divisão do esforço em três grupos de atividades, os quais estavam relacionados ao levantamento do estado da arte e mapeamento de requisitos para certificação digital de documentos e preservação digital distribuída, definição de uma arquitetura genérica para autenticação e preservação de documentos digitais e a montagem de um protótipo para validação da abordagem proposta. Uma visão geral do contexto do serviço e dos principais atores serviço ilustrado na Figura 1.

O serviço proposto não interfere no fluxo natural entre os três atores em pauta mas permite que tanto a atuação do emissor quanto do receptor do documento digital assinado seja facilitada, sobretudo na interação para a validação do mesmo. Sempre centrado no uso do documento digital assinado pelos atores, o serviço provê mecanismos para registrar e preservar o documento para o emissor e para autenticá-lo para o receptor, fazendo ainda a guarda de longa duração do mesmo, o que beneficia todos os atores.

No contexto do piloto, focado em diplomas acadêmicos, o emissor do diploma digital é a IES de Origem, o portador do diploma é o aluno egresso da IES de Origem e o papel do receptor é representado por outras IES, órgãos públicos e empresas privadas para

Figura 1. Visão Geral do Serviço Proposto

as quais o portador do diploma está aplicando para uma posição que exige a apresentação do mesmo.

Acreditamos que com um serviço como o proposto aqui em funcionamento, a autenticidade de documentos (como diplomas ou certificados) emitidos ou revalidados por instituições acadêmicas brasileiras poderá ser mais facilmente verificada, tanto por órgãos públicos quanto por instituições privadas e pessoas físicas.

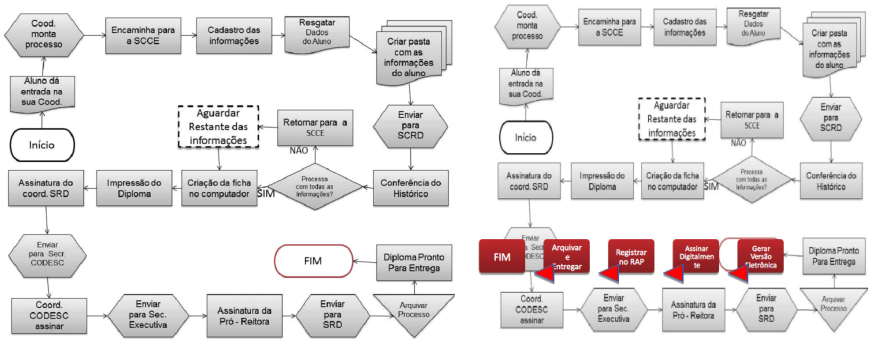
3.1. Estratégia de Modelagem

A estratégia de modelagem do protótipo partiu da premissa de garantir uma integração pouco invasiva com os processos e sistemas em uso atualmente para emissão e registro de diplomas acadêmicos nas IES. Considerando a necessidade de uma transição suave para permitir o aculturamento progressivo com as novas metáforas e tecnologias envolvidas, considerou-se a possibilidade de uma possível convivência dos dois métodos de emissão de diplomas: tradicional em papel e no formato digital.

Neste sentido, alguns requisitos básicos foram estabelecidos para essa primeira fase do protótipo:

- Tentar minimizar a necessidade de intervenção nos fluxos internos dos setores envolvidos na emissão e registro de diplomas das instituições;
- Considerar que o diploma tradicional, em papel, continuará sendo emitido normalmente e que o diploma digital será uma nova opção para o aluno;
- Identificar e aplicar os mesmos protocolos e níveis de alçada usados para a assinatura tradicional dos diplomas em cada instituição na assinatura digital da versão eletrônica;
- Garantir que todo o controle e autonomia presentes hoje nas IES permaneçam inalteradas para a assinatura digital da versão eletrônica. Não deve existir nenhuma transferência de alçada e responsabilidade para o serviço proposto;
- As operações de registro e autenticação de documentos devem ser sempre lastreadas pela validação das assinaturas digitais dos documentos digitais envolvidos e não apenas baseadas na autenticação de usuários e sessões.

Para ilustrar melhor o contexto, a Figura 2 (a) traz o fluxo operacional atualmente



(a) (b)

Figura 2. Fluxo Operacional para Emissão de Diplomas na UFPB (a) e Possível Fluxo de Transição (b)

praticado na UFPB para a emissão de diplomas. Como pode ser visto, são feitas três assinaturas no diploma durante o processo de emissão e por atores/setores diferentes.

Para permitir uma conciliação com tais fluxos internos, que variam de instituição para instituição, a integração com o serviço proposto pode iniciar apenas a partir do momento em que a emissão do diploma tradicional em papel é concluída (Figura 2 (b)). Isso garante que todas as verificações do mérito e das exigências do título acadêmico já foram satisfeitas e concluídas.

A partir deste ponto e contando com o desejo e anuência expressa do interessado, o processo adicional de emissão e registro do diploma digital pode ser iniciado pela própria IES. Esta fase possui cinco etapas lógicas bem definidas:

- Geração de uma versão eletrônica do diploma do aluno;
- Assinatura digital da versão eletrônica pelo(s) mesmo(s) responsável(eis) formal(is) na instituição que assinaram a versão impressa;
- Submissão da versão assinada para registro na DLT e preservação digital no serviço proposto;
- Recepção e arquivamento do recibo de registro na DLT;
- Entrega ao aluno da versão eletrônica assinada do diploma e também do recibo do registro eletrônico.

Estas cinco etapas podem ser implementadas de diferentes formas em cada instituição. Dependendo da maturidade tecnológica e da disponibilidade de recursos, esses passos podem ser integrados ao sistema de controle acadêmico da instituição ou podem ser tratados em uma aplicação específica, por exemplo.

Para permitir a integração, o protótipo de serviço oferece uma interface programável (API) para que instituições de ensino possam registrar (e também validar) diplomas acadêmicos nas DLTs desejadas. Há também uma interface interativa (Portal) para que os interessados possam verificar a autenticidade de um documento através do seu número de registro.

4. Prototipação da Plataforma

4.1. Arquitetura do Protótipo

O protótipo do serviço de registro, autenticação e preservação de documentos acadêmicos utiliza-se de módulos clientes e módulos servidores. Os módulos clientes são utilizados pe-

los usuários para acessar o serviço. Os módulos servidores fornecem os serviços demandados pelos módulos clientes através de APIs públicas. O protótipo permite o uso de diferentes formas de construção e uso de módulos clientes para acessar os serviços oferecidos.

Dessa forma, os módulos clientes usados para as operações de registro e preservação de diplomas podem ser *gateways* para o serviço proposto que operam de forma embutida e/ou integrada aos sistemas de gestão acadêmica das instituições (como o SIGAA, por exemplo) ou podem ser aplicações autônomas, construídas especificamente para esse fim. Os módulos clientes de registro e preservação são, prioritariamente, destinados para uso pelas instituições de ensino credenciadas para a utilização do serviço.

Os módulos clientes exclusivos para autenticação, por sua vez, podem ser portais interativos ou aplicações *desktop* ou *mobile*, e são destinados a usuários finais e instituições que precisam fazer, por razões variadas, a verificação da autenticidade e a prova da existência do registro de um determinado documento digital.

Os módulos clientes de registro e preservação além dos módulos clientes de autenticação tanto podem acessar a API do protótipo diretamente quanto podem fazer uso de componentes que tornam transparente essa integração, chamados aqui de agentes. São previstos dois agentes distintos: i) um que encapsula as operações de registro e preservação, e ii) outro que encapsula as operações de autenticação.

O protótipo propriamente dito é composto por um conjunto de serviços, conforme pode ser visto na Figura 3: *Módulo de Registro*, *Módulo de Autenticação* e Módulo de Preservação. O acesso a tais serviços é intermediado por uma *API RESTful*, chamada *RAP Server*[5]. A API é consumida pelos módulos clientes diretamente ou através do *Agente de Registro* e do *Agente de Autenticação*. Há ainda um *Portal de Autenticação*, o qual permite o acesso interativo para as funcionalidades de autenticação através de um *browser*, e dois componentes internos: o *DLT Broker*, que encapsula a comunicação com as diversas DLTs suportadas, e o *Curador Digital*, *middleware* que cuida da preservação digital.

O *Módulo de Registro* (MR) é o componente do protótipo responsável por providenciar que uma prova de existência do documento digital de entrada seja efetivamente registrada na(s) DLT(s) adequada(s).

Para a seleção de qual (ou quais, se o cliente optar por mais de um) DLT deve ser utilizado pelo MR, são usados parâmetros indicados na própria solicitação ou, na ausência destes, regras e heurísticas pré-definidas pela IES emissora. Embora o escopo do protótipo contemple apenas o registro nas cadeias *Bitcoin* e *Ethereum*, esta generalização considera um ciclo de vida de longa duração para o serviço, no qual novos DLTs concorrentes poderão surgir, oferecendo preços e condições diferenciadas. Neste sentido, é fundamental que o serviço ofereça flexibilidade de escolha para os clientes e suporte para o registro em DLTs diferentes, inclusive permitindo registrar um mesmo documento em mais de um deles.

O MR é invocado pelo *RAP Server* dentro do contexto de uma operação de registro de um novo documento digital e, após um pré-tratamento e validação da legitimidade do documento de entrada, faz a seleção do(s) DLT(s) adequado(s), obedecendo alguma heurística indicada pela IES emissora, e, em seguida, faz o repasse do pedido de registro para o *DLT Broker*, o qual fará o depósito efetivo no(s) DLT(s) indicado(s).

No pré-tratamento do documento digital, além de outras ações, será feita a con-

[5]RAP é um acrônimo para Registro, Autenticação e Preservação.

Figura 3. Arquitetura do Protótipo

ferência da(s) assinatura(s) digital(is) aplicada(s) no documento a ser registrado. Além da corretude da(s) assinatura(s) digital(is) também será aferido se os certificado(s) utilizado(s) estão em conformidade, em quantidade e identidade, com o protocolo e os ids dos certificado(s) cadastrados pela IES emissora para registro do tipo de documento em pauta. Esse passo permite um controle fim-a-fim da legitimidade e integridade dos documentos digitais manuseados pelo serviço e é usado em vários pontos do protótipo.

Após as fases de validação do documento e seleção do(s) DLT(s), o MR submete o pedido de registro devidamente instruído para o *DLT Broker*. Caso o registro seja bem sucedido, o *DLT Broker* devolve o(s) respectivo(s) recibo(s) para o MR, o qual por sua vez, o(s) retorna para o *RAP Server*.

O *Módulo de Autenticação* (MA) é responsável por verificar a autenticidade de diplomas digitais através da recuperação do diploma no repositório de preservação e da validação do registo do referido documento em um dos DLTs distribuídos suportados pelo sistema.

Por meio do módulo de autenticação, um módulo cliente pode fazer uma solicitação de validação de um documento digital. Essa solicitação suporta diferentes formas de validação e recuperação de um documento: i) a partir do recibo da transação registrada em uma DLT; ii) via *upload* de uma cópia do documento digital; iii) pelo número de registro do documento na instituição emissora; iv) pelo *hash* do arquivo, e; v) através de metadados relacionados ao documento ou ao portador/beneficiário do mesmo. O processo de autenticação do MA é invocado a partir do *RAP Server*.

O MA também oferece *endpoints* de acesso para consulta com opções diversificadas de busca ao diploma e/ou registro. Quando uma requisição de consulta chega ao MA, é verificado se o documento desejado existe e então é feito o procedimento de recuperação das informações associadas. Dependendo do tipo de informação de recuperação fornecido, a consulta e recuperação das informações pré-validação é executada em uma ordem específica.

Caso a consulta esteja sendo feita por meio do recibo da transação de registro em uma DLT, o MA irá, inicialmente, se comunicar com o *DLT Broker* a fim de recuperar os

dados de registro na respectiva DLT. Uma vez retornado o registro, a busca pelo documento digital será feito na base de repositório de preservação através do *hash*, o qual representa unicamente o documento digital e que foi extraída dos dados registrados na DLT.

Caso a consulta esteja sendo feita com alguma informação diferente, a recuperação dos dados será iniciada por meio da comunicação entre o MA e o repositório de preservação. Nesse caso, alguns dos filtros de busca que podem ser utilizadas são: o próprio documento digital; um identificador específico relacionado ao documento (da IES emissora ou do próprio serviço); algum dado pessoal do detentor do diploma (como CPF, nome etc) ou algum outro classificador contextual que permita relacionar o diplomado/instituição ao documento, como turma concluinte, ata de colação, livro, página e número de registro etc.

Para garantir a privacidade dos documentos registrados e preservados, alguns dos filtros de recuperação citados são de uso exclusivo da IES emissora. Outras IES, empresas, instituições e demais interessados em validar um diploma na plataforma precisará ter acesso à versão digital do mesmo ou, no mínimo, ao recibo de registro na DLT ou a um dos identificadores únicos. Nestes casos, é esperado que o fornecimento de tais informações de acesso seja feito diretamente pelo próprio detentor do diploma.

O *Módulo de Preservação* (MP) é o componente do protótipo que encaminha cada documento digital registrado no serviço para que este seja preservado e esteja continuamente acessível e utilizável por longo prazo, independentemente da continuidade do sistema utilizado (e/ou da IES emissora) que fez o seu registro.

O MA serve como uma abstração para acesso ao sistema de curadoria digital que é utilizado, e é invocado toda vez uma operação de armazenamento ou recuperação de dados preservados é disparada pelo *RAP Server*.

As operações manipulam pacotes SIP (do inglês *Submission Information Package*), os quais contém, além do documento a ser preservado, metadados com informações extras como o EAD (*Encoded Archival Description*) que descreve o conteúdo do arquivo, junto com informações do autor, data de publicação, entre outras informações a respeito do conteúdo preservado, seguindo as orientações da norma ISO 14721:2003, a qual estabelece o modelod e referência para um *Open Archival Information System* (OAIS). Outros metadados usados pelo MA incluem o recibo de registro gerado pela(s) DLT(s) usadas e também o PREMIS (*Preservation Metadata Maintenance Activity*), que armazena dados sobre o documento de forma a subsidiar futuras ações de preservação do *middleware* de curadoria, como atualização de formato, validação de réplicas e/ou mudança de suporte físico, por exemplo.

4.2. Estudo Preliminar de Integração

Após o desenvolvimento do protótipo e dos agentes de registro e autenticação, o passo seguinte foi fazer uma validação preliminar da flexibilidade e adequação do protocolo de comunicação e da mensageria implementados pelo *RAP Server* e também da funcionalidade dos agentes.

As IES parceiras, UFPB e PUC-RJ, possuem ambientes e *workflows* distintos para controle da emissão de diplomas. A primeira utiliza o SIGAA, o qual já oferece um módulo para o controle de emissão de diplomas. A PUC-RJ, por sua vez, faz a gestão dos concluintes e diplomas usando um sistema específico. A presença de duas realidades bem distintas fez com que os mecanismos de integração previstos no protótipo fossem exigidos e revisados, mesmo com uma amostra pequena de IES.

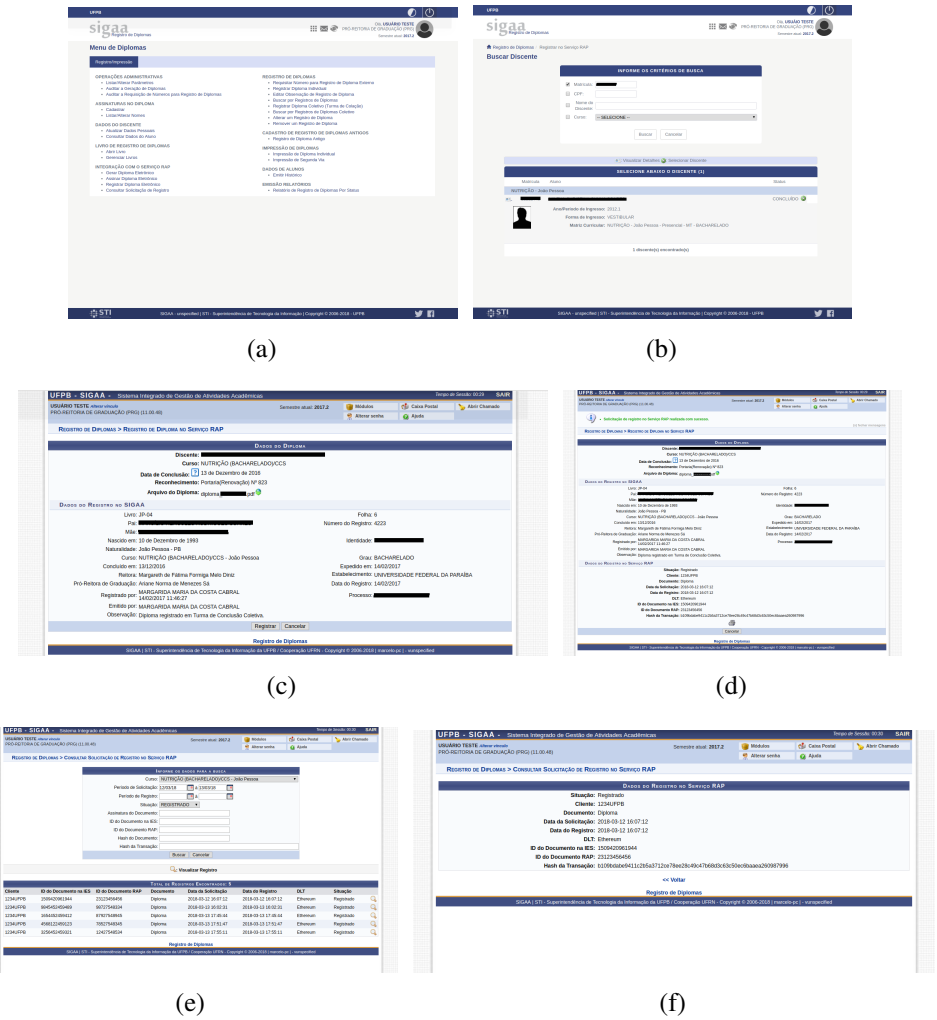


Figura 4. Telas Ilustrativas do Fluxo de Integração SIGAA/UFPB com o Protótipo RAP

No caso da UFPB, o *Módulo de Diplomas* do SIGAA ganhou um submenu específico para integração com o protótipo como pode ser visto na Figura 4 (a). A abordagem usada na UFPB é centrada na emissão individual de diplomas, como ilustrado na Figura 2, e a solicitação de registro e preservação de diplomas é feita diploma a diploma (Figuras 4 (b), (c) e (d)). O *Agente de Registro*, o qual opera localmente, além de fazer a comunicação como o *RAP Server* para o registro de diplomas, também é ativado pelo SIGAA para realizar a geração do documento digital no formato PDF/A e também para assiná-lo digitalmente.

As solicitações de registro pendentes são acompanhadas por pedidos de status ao *RAP Server* e, quando efetivadas, são sinalizadas na base local do SIGAA, permitindo que tanto o documento digital quanto o recibo de registro na DLT sejam armazenados junto ao cadastro do egresso, assim como entregues ao interessado. As Figuras 4 (e) e (f) mostram as telas de pesquisa e recuperação da informação de registro de diplomas na DLT.

A PUC-RJ, por sua vez, possui um sistema específico para controle e emissão de diplomas e a forma de integração adotada pela equipe local foi a construção de um *gateway* para fazer a interface com o *RAP Server*. Uma outra diferença com relação à UFPB foi a abordagem baseada em turmas de formandos, tratando o registro de diplomas como

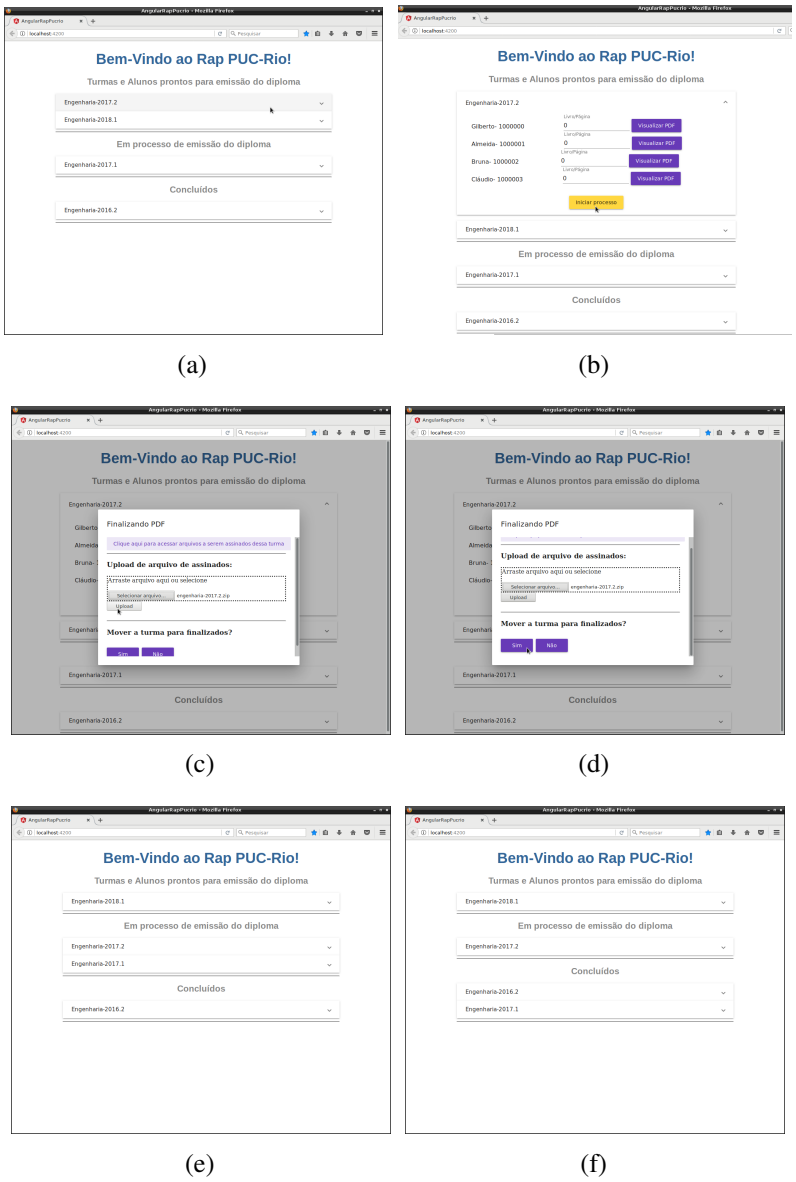


Figura 5. Telas Ilustrativas do Fluxo de Integração da PUC-RJ com o Protótipo RAP

operações em lote.

As Figuras 5 (a-f) ilustram o *workflow* de preparação das turmas para registro, gerando e carregando os diplomas individualmente para, só então, disparar as solicitações de registro uma a uma. Apenas quando todos os diplomas de uma turma têm os seus registros confirmados é que a turma é marcada como finalizada.

Essa é uma estratégia interessante que motivou a realização de um ajuste no protótipo para suportar a abertura e fechamento de lotes de solicitações de registro, assim como a realização de registros múltiplos na DLT através mecanismos como Árvores de Merkle para diminuir os custos, no caso de uso de DLTs tarifadas.

5. Conclusão

A principal contribuição científica deste projeto é a investigação de uma nova abordagem baseada no uso de DLTs e repositórios ativos distribuídos para a autenticação e preservação

de longo prazo de documentos digitais legais.

Do ponto de vista tecnológico, um dos principais resultados deste trabalho foi a criação de um ambiente para dar suporte, inicialmente, a um protótipo de Serviço para Autenticação e Preservação de Documentos Digitais Acadêmicos, com potencial para adoção por diversas organizações educacionais, sejam públicas ou privadas.

Este estudo de caso específico de autenticação digital apresenta um excelente potencial de aplicação real não financeira de DLTs e pode ser o embrião para a oferta futura de um serviço permanente de grande utilidade para as IES nacionais, parcela significativa da comunidade de usuários da RNP.

O serviço proposto oferece também mecanismos que possibilitem sua integração futura com diferentes plataformas que fazem uso deste tipo de documento, como a Plataforma Lattes do CNPQ, utilizada para o cadastro de pesquisadores no país, o sistema e-MEC, utilizado pelo INEP nos processos de avaliação do ensino superior no Brasil, a plataforma Sucupira, utilizada pela CAPES no processos de avaliação da pós-graduação no Brasil e o SIGAA, sistema de controle acadêmico desenvolvido pela UFRN e amplamente utilizado por Instituições Federais de Ensino Superior.

A integração do serviço proposto com as plataformas mencionadas pode melhorar o processo de cadastramento de informações profissionais nestes sistemas, uma vez que permitirá a validação automática da autenticidade dos diplomas e certificados fornecidos, adicionando eficiência e economia ao processo e ajudando a impedir que situações como as apresentadas anteriormente voltem a acontecer.

Porém, cabe ressaltar que a infraestrutura a ser desenvolvida para dar suporte ao serviço proposto poderá apoiar também a criação de vários outros serviços para preservação e autenticação de documentos digitais e não apenas de diplomas.

Do ponto de vista da preservação digital e por conta das características de encapsulamento da replicação e do controle de falhas que pretendemos inserir na camada de armazenamento, tal abordagem pode trazer uma série de resultados complementares:

- Avançar na direção de uma integração transparente da funcionalidade de repositórios ativos com as outras camadas do modelo OAIS [Lavoie 2000];
- Habilitar a oferta de níveis distintos de capacidade de preservação, de acordo com a relevância do objeto digital;
- Possibilidade de operar em diferentes contextos de disponibilidade e capacidade de recursos computacionais.

6. Agradecimentos

Os autores gostariam de agradecer a RNP pelo financiamento desta pesquisa através do programa de Grupos de Trabalho (GTs) e também as equipes da Superintendência de Tecnologia da Informação (STI) da UFPB e da Divisão de Admissão e Registro (DAR) e do Laboratório Telemídia da PUC-RJ pela inestimável ajuda no fornecimento de requisitos e na adaptação dos respectivos sistemas para integração com o protótipo RAP.

Referências

Baguete (2013). Certisign oferece diploma digital. http://www.baguete.com.br/noticias/22/02/2013/certisign-oferece-diploma-digital. [Online; accessed 21-March-2018].

Costa, R., Lemos, G., Becker, V., and Malaguti, A. (2015). Estratégias para criação de de uma rede nacional para preservação digital de acervos audivisual brasileiros. *Reflexões sobre Preservação Audiovisual/10 anos da CineOP – Mostra de Cinema de Ouro Preto.*

Costa, R., Lemos, G., Becker, V., and Malaguti, A. (2016). We need to talk about digital preservation of audiovisual collections: Strategies for building national networks. *jAUTI 2016: V Iberoamerican Conference on Applications and Usability of Interactive TV / 18 Convención Científica de Ingineria Y Arquitectura.*

Crosby, M., Pattanayak, P., Verma, S., and Kalyanaraman, V. (2016). Blockchain technology: Beyond bitcoin. *Applied Innovation*, 2:6–10.

do Povo, G. (2013). Mp investiga o uso de mais de 500 diplomas falsos em maringá.

Estadao (2016). Sem papel, usp atrasa impressão de diplomas por quatro meses. http://educacao.estadao.com.br/noticias/geral,sem-papel-usp-atrasa-diplomas-por-quatro-meses,10000064951. [Online; accessed 21-March-2018].

Ferreira, M., Saraiva, R., and Rodrigues, E. (2012). Estado da arte em preservação digital.

FOLHA (2011). INEP cancela avaliação de quatro cursos por suspeita de fraude. http://www1.folha.uol.com.br/saber/941107-inep-cancela-avaliacao-de-quatro-cursos-por-suspeita-de-fraude.shtml. [Online; accessed 21-March-2018].

GLOBO (2013). Esquema de revalidação de diploma de medicina é desarticulado pela pf.

GLOBO (2017). Dono de universidade denuncia esquema de venda de diplomas falsos por r$ 550,00 em mt.

ITI (2013). Usp adota diploma com certificação digital. http://www.iti.gov.br /noticias/indice-de-noticias/4230-usp-adota-diploma-com-certificacao-digital. [Online; accessed 21-March-2018].

Kakavand, H., Kost De Sevres, N., and Chilton, B. (2017). The blockchain revolution: An analysis of regulation and technology related to distributed ledger technologies.

Lavoie, B. (2000). Meeting the challenges of digital preservation: The oais reference model. *OClC Newsletter*, 243:26–30.

Online, G. (2017). Sedu pode levar até 180 dias para investigar diplomas falsos - mais de 100 professores processados por usarem diplomas falsos no es.

Ruusalepp, R. and Dobreva, M. (2012). Digital preservation services: state of the art analysis. technical report, dc-net.

Sayão, L. F. and Sales, L. F. (2012). Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. *Informação & Sociedade*, 22(3).

Skinner, K. and Halbert, M. (2009). The metaarchive cooperative: a collaborative approach to distributed digital preservation. *Library Trends*, 57(3):371–392.

Skinner, K. and Schultz, M. (2010). *A guide to distributed digital preservation*. Lulu. com.

Wikipedia (2018). Proof of existence. https://en.wikipedia.org/wiki/Proof_of_Existence. [Online; accessed 21-March-2018].