

Prémio Arquivo.pt

Descrição Sumária do Trabalho

Identificação

- Título: ArchiveChain – Arquivar o passado, impulsionar o futuro.
- Área temática: Blockchains, Sustentabilidade energética e ambiental, Preservação Digital da Web
- Candidato: Bruno Cotrim, Bernardo Ferreira, Miguel Matos
- Email: fc54406@alunos.ciencias.ulisboa.pt, blferreira@ciencias.ulisboa.pt, miguel.marques.matos@tecnico.ulisboa.pt,

Descrição do Trabalho

A ArchiveChain é uma nova blockchain sustentável que utiliza as páginas do Arquivo.pt para evitar os desperdícios energéticos e requisitos de hardware especializado de blockchain tradicionais.

Atualmente, existem três grandes tipos de blockchains públicas: (i) as baseadas em provas de trabalho (ex: Bitcoin), que usam poder computacional para decidir que blocos e transações adicionar à blockchain, consumindo quantidades maciças de energia; (ii) as baseadas em provas de participação (ex: Ethereum), que resolvem o problema energético das anteriores mas promovem que quem tenha mais fundos (a participação) tenha maior poder de decisão e acumule cada vez mais sem qualquer contrapartida, tornando-se muito difícil para novos participantes competirem; (iii) e as baseadas em provas de espaço (ex: Chia), que resolvem o problema energético trocando o uso de poder computacional por armazenamento de dados em disco, no entanto os dados armazenados são inúteis (bytes aleatórios) e servem apenas para as ditas provas, levando a um grande desperdício de recursos de armazenamento.

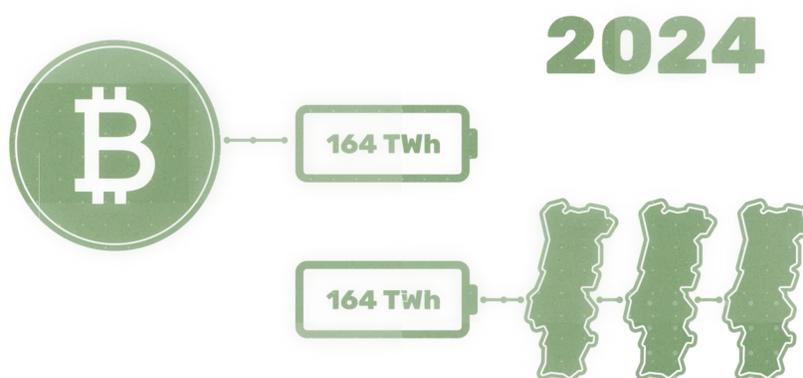


Figura 1 – Em 2024, a Bitcoin consumiu 3x mais energia que Portugal.

A ArchiveChain resolve estes problemas propondo o conceito de provas de espaço útil, onde o espaço em disco dos participantes, em vez de usar bytes aleatórios, utiliza as páginas do Arquivo.pt, juntamente com garantias criptográficas de que estas são armazenadas de forma intacta. Desta forma a ArchiveChain consegue atingir um consumo energético 1000 vezes inferior ao da Bitcoin, enquanto resolve o problema de espaço inútil da Chia e evita o acúmulo sem contrapartida de fundos e poder de decisão da Ethereum, conseguindo assim ser uma solução mais sustentável, democrática e justa para implementar cripto-moedas e contratos inteligentes.

Simultaneamente, ao utilizar as páginas do Arquivo.pt, a ArchiveChain promove e ajuda na missão de arquivar a web portuguesa. Os participantes da ArchiveChain, ao entrarem na rede, recebem automaticamente páginas do Arquivo.pt para usarem nas suas provas de espaço útil. Quem armazenar mais páginas, tem maior chance de decidir os próximos blocos e receber recompensas em forma de cripto-moeda. Periodicamente, são também desafiados a provar a posse e integridade das páginas que armazenam, garantindo assim a sua disponibilidade e correto armazenamento. Por fim, quando um utilizador do Arquivo.pt quer consultar uma página histórica do arquivo, esta pode ser obtida diretamente de um dos participantes da ArchiveChain, ajudando assim a reduzir os custos operacionais do Arquivo.pt. Este serviço prestado pela ArchiveChain pode ser implementado de forma completamente transparente para o utilizador final, mantendo o Arquivo.pt jurisdição completa e controlo de qualidade de que páginas são arquivadas e por quem.

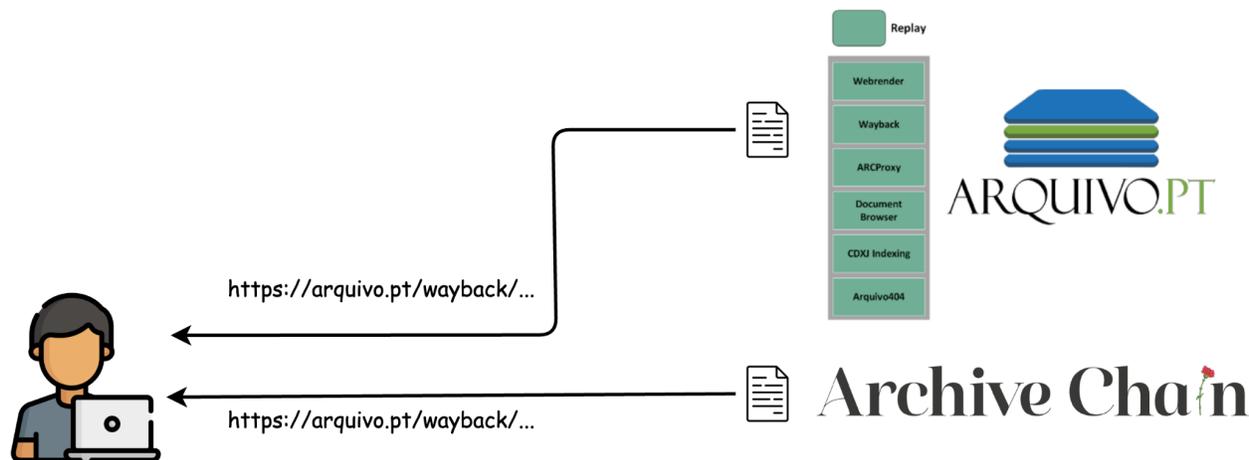


Figura 2 – A ArchiveChain como suporte à base de dados atual do Arquivo.pt.

A ArchiveChain permite que todos os cidadãos possam ter um papel ativo no arquivamento de páginas da web portuguesa, transformando a preservação do património digital numa missão colaborativa, acessível e sustentável. Ao participar na ArchiveChain, quem contribuir com espaço em disco para armazenar páginas do Arquivo.pt recebe incentivos proporcionais à quantidade de espaço alocado e número de páginas armazenadas, incentivando os participantes a armazenar o maior número de páginas possíveis para maximizar o seu retorno e podendo assim

também dar uma nova vida ao seu espaço em disco que de outra forma não estaria a ser aproveitado.

Total Stored (TB) per Year Cumulative

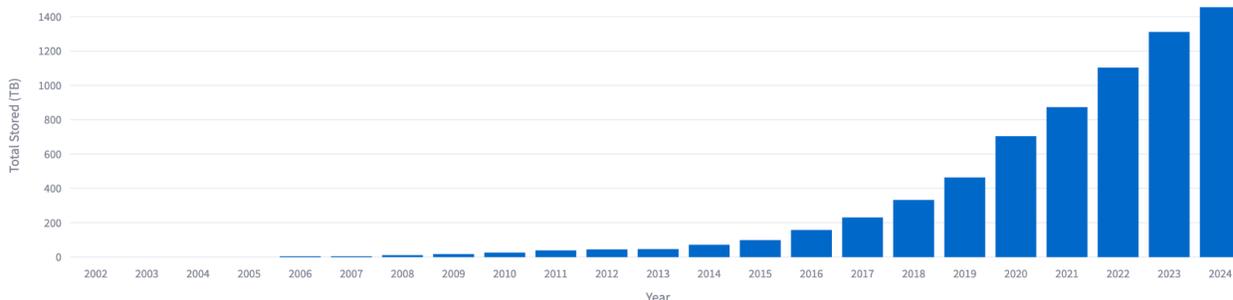


Figura 3 - Evolução da quantidade de dados acumulados pelo Arquivo.pt ao longo dos anos.

Objetivos

Os principais objetivos da ArchiveChain são:

- Democratizar a preservação do património digital, permitindo aos cidadãos contribuir ativamente para o arquivamento da web portuguesa, com incentivos na forma de cripto-moedas.
- Proporcionar uma solução escalável que acompanhe o crescimento exponencial dos dados do Arquivo.pt, permitindo uma redução dos seus custos de operação por meio da descentralização e terceirização do armazenamento, com transparência total para o utilizador final do Arquivo.pt e sem comprometer a sua qualidade de serviço.
- Construir uma blockchain sustentável e ecologicamente viável, baseada no armazenamento das páginas do Arquivo.pt, que permita a implementação de cripto-moedas e contratos inteligentes.

Resultados Atingidos

Os seguintes resultados foram atingidos com este projeto:

Blockchain Sustentável: Foi implementada uma blockchain energeticamente eficiente que utiliza espaço útil em disco como recurso principal, nomeadamente as páginas do Arquivo.pt. Este sistema permite o armazenamento descentralizado das páginas do Arquivo.pt, garantindo integridade e disponibilidade das mesmas. As páginas arquivadas desempenham um papel central no funcionamento da blockchain, conferindo um duplo propósito ao espaço em disco:

preservar o património digital e manter a segurança da blockchain. Um modelo de incentivos baseado em cripto-moedas recompensa os participantes proporcionalmente ao volume de páginas armazenadas, promovendo a participação ativa e reduzindo significativamente os custos operacionais do Arquivo.pt. A integração com a base de dados atual do Arquivo.pt é transparente, mantendo este o controlo total do processo de arquivamento e uma experiência transparente e inalterada para o utilizador final.

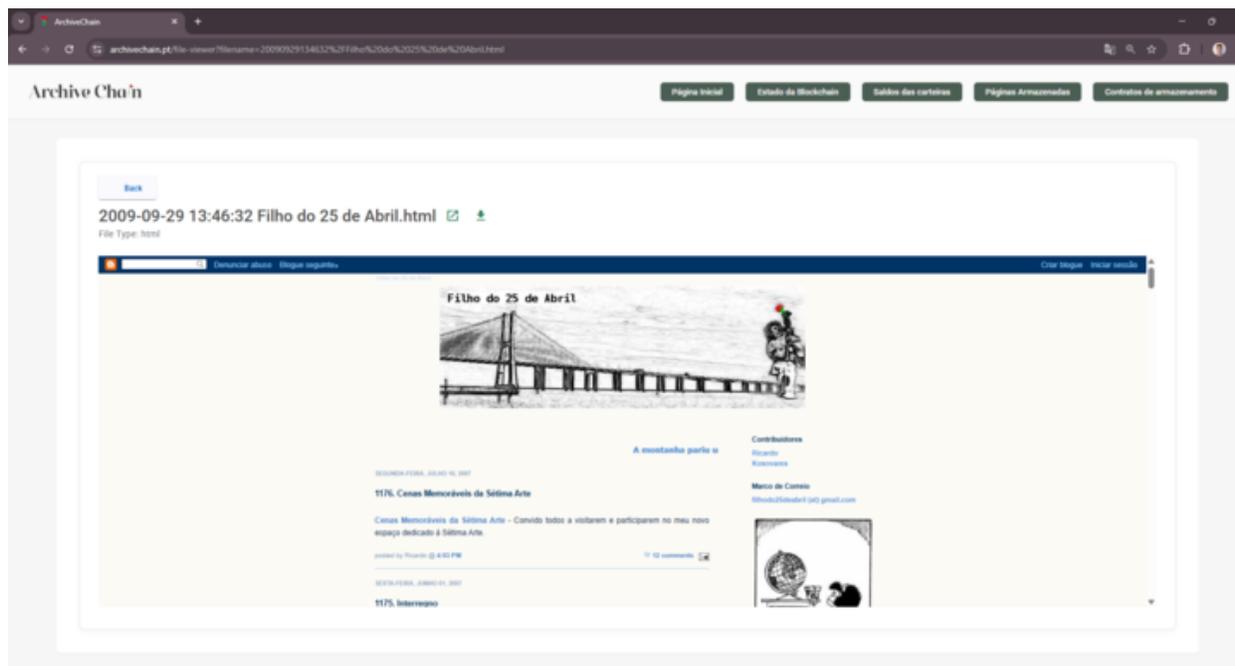


Figura 4 – Página do Arquivo.pt a ser consultada através da plataforma web da ArchiveChain.

Plataforma Web: Foi também desenvolvida uma plataforma web (archivechain.pt) para alojar e demonstrar o projeto. Esta inclui um explorador da blockchain, onde é possível observar o estado atual da ArchiveChain, a criação de blocos em tempo real e métricas globais do sistema através de gráficos interativos. O explorador permite ainda pesquisar as páginas do Arquivo.pt atualmente replicadas pela ArchiveChain e consultar, para cada participante, as páginas arquivadas por este, moedas ganhas, e a sua atividade no sistema.

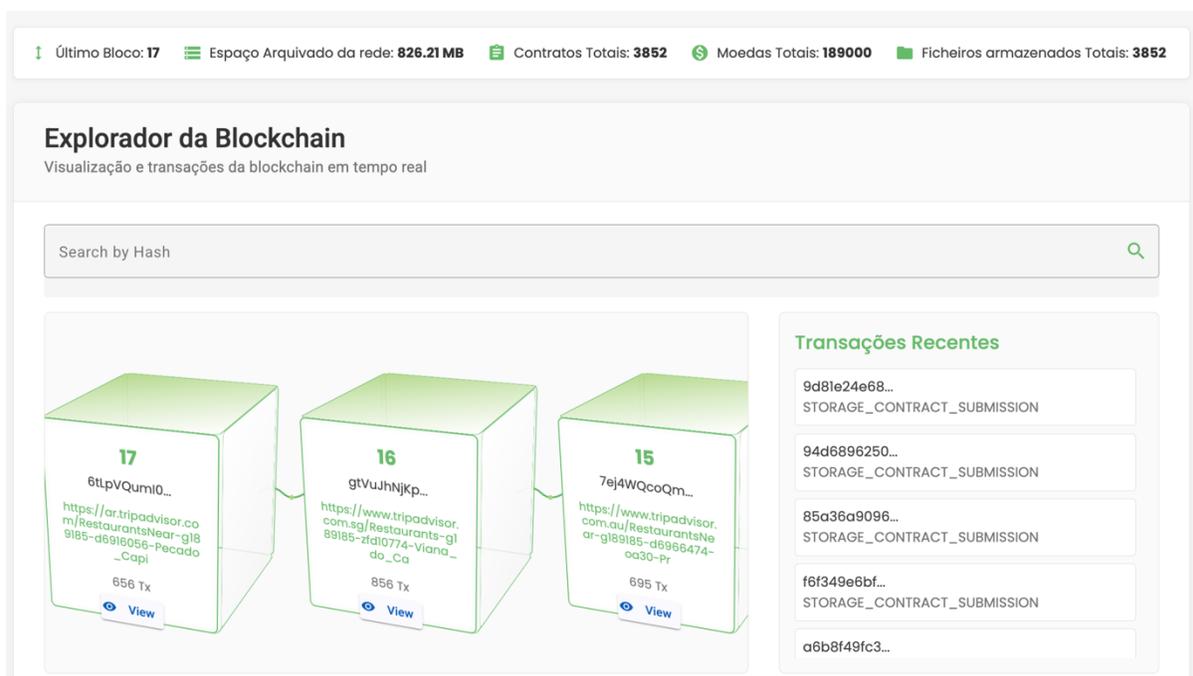


Figura 5 – Visualização da criação de blocos em tempo real e métricas globais do sistema, consultáveis através da plataforma web da ArchiveChain.

Originalidade e caráter inovador

A ArchiveChain apresenta uma abordagem inovadora, sendo do nosso conhecimento a primeira blockchain do mundo que utiliza dados de arquivos web como provas de espaço útil. Em particular, a ArchiveChain utiliza as páginas do Arquivo.pt como parte do seu funcionamento, permitindo assim a manutenção segura da blockchain e resolvendo alguns dos problemas das blockchains tradicionais, como o consumo massivo de energia e o desperdício de espaço de armazenamento digital.

Simultaneamente, a ArchiveChain incentiva os cidadãos a fazerem parte da missão coletiva de arquivar o nosso passado, armazenando as páginas históricas da web portuguesa, protegendo-as contra perdas, e disponibilizando-as para consulta, cooperando assim com o Arquivo.pt nesta missão. Ao recompensar os seus participantes com cripto-moedas, a ArchiveChain alinha incentivos, promove o interesse dos cidadãos na missão de arquivar a nossa história digital, e torna-se autossustentável.

Em contraste com trabalhos anteriores, focados principalmente em recolher dados do Arquivo.pt para coleta e processamento de informação, a ArchiveChain demonstra uma abordagem única focada no armazenamento distribuído e replicação das páginas arquivadas, integrando-se com as bases de dados atuais do Arquivo.pt de forma transparente e ajudando a resolver o problema de escalabilidade de armazenamento e redução de custos operacionais resultantes do aumento exponencial de conteúdos digitais ao longo dos anos. A ArchiveChain oferece assim uma

metodologia completamente diferente de utilização e oferta de valor de serviço do Arquivo.pt, focando em problemas e soluções nunca antes explorados.

Impacto social (aplicação e utilidade social)

Na era digital, preservar o passado é tão vital quanto avançar para o futuro. A ArchiveChain abraça ambas as missões, usando páginas do Arquivo.pt para construir uma blockchain que consome mil vezes menos energia que a Bitcoin, cuja pegada ecológica de 164 TWh triplica o consumo energético português. Num mundo onde a energia se tem tornado um luxo escasso, famílias e serviços públicos enfrentam dificuldades diárias. A ArchiveChain devolve esse recurso à sociedade, promovendo um planeta mais verde e tecnológico. Ao mesmo tempo, a ArchiveChain cria um sentido de comunidade onde qualquer pessoa pode fazer parte da missão de arquivar o nosso passado digital, protegendo-o contra perdas, garantindo a sua disponibilidade, e ganhando cripto-moedas como recompensa. Por fim, a ArchiveChain inspira novos projetos que priorizem a sustentabilidade ecológica e eleva a literacia tecnológica, desmistificando inovações complexas e unindo sustentabilidade e preservação digital num ecossistema único.

Impacto científico (aplicação e utilidade científica)

A ArchiveChain é a primeira blockchain do mundo que usa dados de arquivos web no seu funcionamento interno. Para tal, combina técnicas inovadoras de provas de espaço útil (*Proof-of-Useful-Space* - PoUSp) e provas de posse de dados (*Proof-of-Data-Possession* - PDP), utilizando as páginas do Arquivo.pt como base. As PoUSp permitem o funcionamento interno da blockchain utilizando espaço útil em disco como primitiva energeticamente eficiente e as PDP garantem integridade e verificação contínua dos ficheiros armazenados pelos participantes, permitindo assim o armazenamento descentralizado e correto das páginas do Arquivo.pt.

As técnicas desenvolvidas têm valor científico independente e planeamos no médio prazo submeter este trabalho para publicação científica em fórum internacional. Estas técnicas permitirão à comunidade científica internacional estudar novas formas de funcionamento seguro e sustentável para blockchains, oferecendo uma alternativa mais eficiente e aproveitadora de recursos em comparação com abordagens tradicionais.

Além disso, a ArchiveChain introduz um modelo económico baseado em incentivos, recompensando os participantes com cripto-moedas por armazenar as páginas do Arquivo.pt, por as servir e por manter a rede. Este aspeto é crucial para promover a adoção e viabilidade de blockchains em larga escala, como sistemas financeiros e de armazenamento distribuído que permitem implementar contratos inteligentes, analisando como incentivos influenciam a participação e estabilidade do sistema e a sua auto-sustentabilidade.

A interface web e API da ArchiveChain oferecem acesso completo aos seus dados interno, permitindo uma fácil compreensão e análise científica do sistema. Isto permite que outros investigadores possam analisar independentemente vários aspetos da blockchain, tais como o equilíbrio do sistema de incentivos, a distribuição do espaço de armazenamento das páginas do Arquivo.pt, e outras características.

Relevância da utilização do Arquivo.pt

O Arquivo.pt desempenha um papel fundamental na ArchiveChain, fornecendo as páginas web que são usadas como peças essenciais para garantir a sua segurança. Os participantes da ArchiveChain preenchem o seu espaço em disco com provas de espaço útil, construídas a partir das páginas do Arquivo.pt. Como tal, o armazenamento das páginas pelos participantes serve como prova de alocação de um recurso limitado que é essencial para a evolução segura da blockchain.

Archive Chain

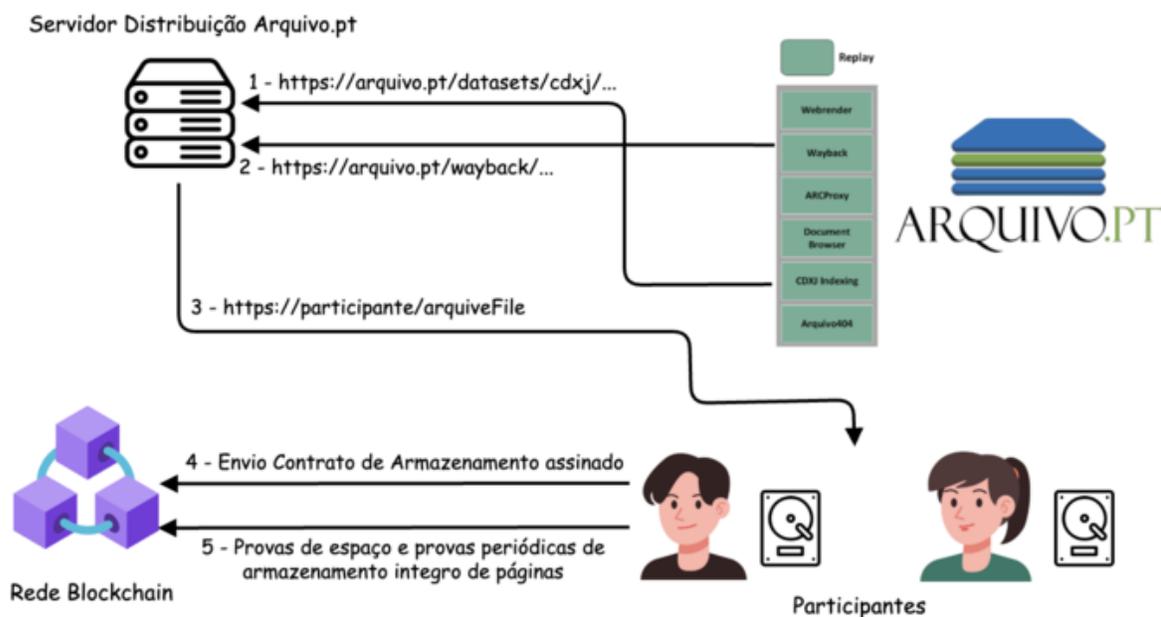


Figura 6 – Ciclo de vida do armazenamento de dados na Archivechain.

Este processo começa com a obtenção de páginas web do Arquivo.pt, utilizando a API CDX para consultar e extrair metadados e URLs de snapshots arquivados. Posteriormente, a API noFrame, acessível através de <https://arquivo.pt/noFrame/replay>, é usada para recuperar os conteúdos originais dessas páginas. As páginas obtidas são então distribuídas pelos participantes da ArchiveChain, onde são pré-processadas e transformadas em provas de espaço útil. Após essa etapa, um contrato de armazenamento, assinado por um servidor do Arquivo.pt (simulado no

protótipo atual) e pelo participante, é registado na blockchain. Este contrato contém informações essenciais para verificar a integridade do armazenamento das páginas pelo participante. Periodicamente, os participantes são desafiados a apresentar provas criptográficas que confirmam a posse e a integridade dos ficheiros armazenados, enviando essas provas à blockchain para validação.

Quando um utilizador solicita uma página, o servidor do Arquivo.pt redireciona o pedido a um participante, que extrai o ficheiro original das provas de espaço e entrega-o ao utilizador. Este processo é transparente, integrando-se facilmente com a base de dados atual do Arquivo.pt, de modo a que o acesso aos ficheiros ocorra sem alterações na experiência do utilizador.

Comentários adicionais

Trabalho realizado no contexto da dissertação de mestrado em Engenharia Informática de Bruno Cotrim, pela Faculdade de Ciências da Universidade de Lisboa, com orientação científica de Bernardo Ferreira (Faculdade de Ciências) e de Miguel Matos (Instituto Superior Técnico).

A equipa agradece ainda o apoio prestado pela Direção da Faculdade de Ciências e pela Direção de Comunicação e Imagem da mesma, em particular a Celma Padamo e a Luís Carriço. Agradecemos também à FCCN pela disponibilização pública e gratuita do Arquivo.pt e da sua API, que foi essencial para concretizar este projeto.

Recursos complementares

- Website Projeto:
 - <https://archivechain.pt/>
- Repositórios Projeto:
 - <https://github.com/brunocotrim1/ArchiveChain>
 - <https://github.com/brunocotrim1/ArchiveChainFrontend>
 - <https://github.com/brunocotrim1/ArchiveChainFCCN>
- Relatório Preliminar Científico:
https://github.com/brunocotrim1/ArchiveChain/blob/main/relatorio_preliminar.pdf