

(DRAFT)

Open VLF, Plataforma de Dados Abertos para *Very Low Frequency Data*

Daniel H. Venna Kauffmann, Lucas Santiago Garcia, Rogério De Oliveira

Faculdade de Computação e Informática (FCI)

Universidade Presbiteriana Mackenzie (UPM) – São Paulo – SP – Brasil

`_@mackenzista.com.br, _@mackenzista.com.br, rogerio.oliveira@mackenzie.br`

Resumo. *Este projeto implementa uma plataforma de dados abertos, flexível, expansível e com suporte a múltiplos repositórios de dados, para dados das antenas da rede SAVNET do CRAAM (Centro de Rádio Astronomia e Astrofísica do Mackenzie). A SAVNET é uma rede de radio antenas de sinais VLF (Very Low Frequency) com onze estações espalhadas na América do Sul, incluindo Brasil, Peru, Argentina e Antártica. São disponibilizados dados coletados desde 2005, com suporte à busca, visualização e download dos dados, incluindo ferramentas de código aberto e conteúdo informativo aos interessados. Assim, a plataforma dará acesso à comunidade científica, incluindo pesquisadores e estudantes, e ao público em geral, a dados importantes para pesquisas sobre clima espacial, erupções solares, e fenômenos de efeitos atmosféricos e eletromagnéticos da ionosfera, com um impacto positivo para pesquisa e o ensino na área.*

Palavras-chave: *Ciência aberta; dados abertos; very low frequency data; VLF; SAVNET.*

1. INTRODUÇÃO

Há hoje um grande movimento para promover a ciência aberta, estimulada por pesquisadores, governos, organizações mundiais e agências de fomento como a FAPESP. Segundo a Unesco (Unesco, 2021), Ciência Aberta envolve vários pilares que vão das publicações científicas abertas e infraestrutura, aos dados de pesquisa abertos (Figura 1), e proporcionam aumento da colaboração científica e do compartilhamento de informações para o benefício da ciência e da sociedade. Este estudo traz uma contribuição para o Conhecimento Científico Aberto, e pretende implementar uma solução de dados de pesquisa abertos que, inicialmente, irá disponibilizar dados de ondas de rádio VLF (Very Low Frequency) da rede de antenas SAVNET (Raulin et al., 2009) do Centro de Rádio Astronomia do Mackenzie (CRAAM) à comunidade científica, incluindo pesquisadores e estudantes, e ao público em geral.

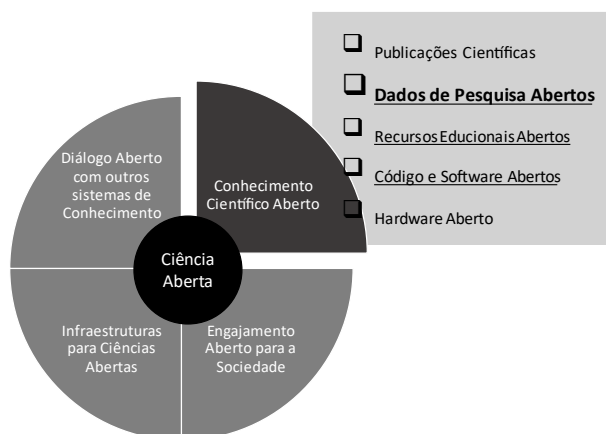


Figura 1. Pilares da Ciência Aberta segundo a UNESCO (Unesco, 2021).

Este trabalho traz uma contribuição no campo de Dados de Pesquisa Aberta, e também para os Recursos Educacionais, Código e Software Abertos.

1.1 Contextualização e Relevância do Tema

Os sinais de VLF, são ondas de rádio frequência muito baixa (3-30 kHz). Eles são usados para serviços de navegação, comunicação com submarinos e são uma ferramenta poderosa para estudar a ionosfera da Terra a baixa altitude (regiões C/D com altitude de 60-70 km) fenômenos atmosféricos e geofísicos,

clima espacial, campo magnético e erupções solares. A SAVNET (Raulin et al., 2009), criada e mantida pelo CRAAM e o INPE (Instituto Nacional de Pesquisas Espaciais), consiste em uma rede de recebimento/rastreamento de 11 estações VLF, localizadas na América Latina e na Antártida. Cinco estações estão no Brasil, 3 no Peru, 1 na Argentina, 1 no México e 1 na Estação de Pesquisa Antártica Brasileira Comandante Ferraz.

O objetivo da rede SAVNET é fornecer dados para uma melhor compreensão da estrutura espacial da Anomalia Magnética do Atlântico Sul (SAMA) (Raulin et al., 2009), o estudo do clima espacial e erupções solares (Raulin et al., 2010), e fenômenos de efeitos atmosféricos e eletromagnéticos (Correia, 2011; Bertoni, 2013; Tacza, 2014). Seus dados incluem sinais diários estreitos (um transmissor) e banda larga (multitransmissores) desde 2005 até agora, em um total de mais de 15 anos dados e 4TB, armazenados localmente (HDs, DVDs etc.) sem qualquer sistema hoje que facilite a busca e descoberta, a consulta ou acesso a esses dados.

1.2 Objeto de Pesquisa

1.2.1 Contextualização do Problema de Pesquisa

A forma de armazenamento dos dados da SAVNET hoje não permite a busca ou o acesso de forma livre e autônoma por pesquisadores ou interessados nesses dados. Mesmo para os pesquisadores locais, com acesso local aos dispositivos de armazenamento, a busca por informações é trabalhosa na falta de sistema de consulta que permite a busca e descoberta dos dados disponíveis. Esta pesquisa busca trazer uma alternativa de acesso a esses dados, que seja aberta e facilmente acessível.

1.2.2 Hipótese

A hipótese deste trabalho é de que é possível desenvolver um sistema aberto de consulta e acesso aos diferentes dados da SAVNET. De que esse sistema pode, assim, contribuir para promover a pesquisa, colaboração e educação, e de que poderá também servir de modelo ou plataforma para o compartilhamento de outros tipos de dados do CRAAM.

1.3 Objetivos do Estudo

1.3.1 Objetivo Geral

O objetivo deste projeto é fornecer uma plataforma de dados abertos, para os dados das antenas VLF da rede SAVNET do CRAAM, contribuindo para promover a pesquisa e colaboração, bem como ações de educação e divulgação científica.

1.3.2 Objetivos Específicos

A solução de dados abertos ainda pretende:

1. Ser desenhada para suportar diferentes repositórios de dados, podendo empregar, ou ser expandida, para acessar armazenados em diferentes tecnologias ou plataformas (sistemas de arquivos locais, diferentes tipos de armazenamento em nuvem etc.).
2. Permitir a inclusão de outros repositórios de dados VLF além dos dados da SAVNET, mesmo que pertencentes a outras organizações, mas desde que seguindo os padrões de dados VLF suportados (AWESOME e SAVNET), sendo aberta não somente para consulta, mas para adição de novos dados tornados públicos.
3. Proporcionar ferramentas que facilitem a busca e o acesso às informações, incluindo ferramentas de pesquisa (*search*), navegação, visualização dos dados e scripts de manipulação dos dados em código aberto.
4. Incluir informações de divulgação científica, democratizando e permitindo o seu acesso a um público mais amplo.

Por último, será idealizada tendo em vista seu reuso ou adaptação para o compartilhamento de outros tipos de dados de pesquisa e equipamentos do CRAAM, como dados de antenas GPS, estações meteorológicas etc.

1.4 Justificativa

Os dados VLF da rede de antenas SAVNET (CRAAM) fornecem, por sua localização na América Latina, dados únicos para pesquisas envolvendo fenômenos da Anomalia Magnética do Atlântico Sul, clima espacial, manchas solares e outros fenômenos atmosféricos e geofísicos. O acesso a esses dados, entretanto, tem uma série de limitações sendo armazenados de forma local e sem ferramentas de busca ou consulta que facilitem o seu acesso. Nina (2020) destaca a importância de grandes bancos de dados relacionados a dados de amplitude e fase de ondas VLF tanto para fins de pesquisa (como o estudo de fenômenos astro e geofísicos, desastres naturais como terremotos etc.) como para aplicações práticas, relacionadas a interferência nas telecomunicações (como comunicações e o sistema global de navegação via satélite) causada por eventos na ionosfera. Deste modo, parece relevante a criação de uma plataforma de dados abertos que, além de fornecer ferramentas para o fácil acesso pelos pesquisadores locais, permita o acesso aberto a esses dados para toda comunidade científica e o público interessado, facilitando, promovendo e divulgando as pesquisas científicas na área.

Há de se destacar que, desde 2017, a FAPESP lançou as bases para a abertura de dados como uma exigência submissão de projetos com os Planos de Gestão de Dados (FAPESP, 2017). Isso vem se tornando uma exigência padrão por outras agências e torna estudos como este, um desenvolvimento quase obrigatório para projetos que requerem auxílio das agências de fomento. Para a FAPESP (FAPESP, n.d.) os *“resultados de pesquisa financiada por verbas públicas são um bem público e como tal devem ser disseminados em larga escala, resguardados os aspectos éticos, de privacidade, segurança e relativos à proteção de propriedade intelectual”*, sendo a disseminação alcançada através dos instrumentos da ciência, dados e códigos abertos.

1.5 Organização do Estudo

Este projeto está organizado conforme os capítulos descritos a seguir:

As demais seções deste trabalho encontram-se organizadas do seguinte modo. O capítulo 2, Referencial Teórico, apresenta os principais conceitos que envolvem

este trabalho. O capítulo 3, Metodologia da Pesquisa, discute a metodologia a ser empregada na realização deste estudo. Ao final, o capítulo 4, Cronograma, apresenta as fases e datas previstas para realização deste estudo.

2. REFERENCIAL TEÓRICO

Para este estudo são particularmente importantes conceitos sobre ciência e dados abertos, que motivam este projeto, e sobre VLF, e suas aplicações no campo da pesquisa. Esses conceitos são brevemente desenvolvidos aqui.

2.1. Ciência e Dados Abertos

A ciência aberta é um movimento para tornar a pesquisa financiada publicamente mais amplamente acessível à comunidade científica, às empresas e a sociedade em geral (OCDE, 2015). Ela envolve iniciativas que vão da publicação de trabalhos científica ao uso de dados abertos, passando por infraestruturas de suporte, código aberto etc. (Unesco, 2017). Há diferentes definições de Ciência Aberta. A Comissão Europeia (EC, 2021), por exemplo, define a ciência aberta como *"uma nova abordagem do processo científico baseada no trabalho cooperativo e em novas formas de difundir o conhecimento através da utilização de tecnologias digitais e de novas ferramentas colaborativas"*. Foster (n.d.) define a ciência aberta como *"...a prática da ciência de tal forma que outros possam colaborar e contribuir, onde os dados de pesquisa, notas de laboratório e outros processos de pesquisa estão disponíveis gratuitamente, sob termos que permitem a reutilização, redistribuição e reprodução da pesquisa e seus dados e métodos subjacentes"*, e apresenta ainda uma taxonomia completa de termos relacionados como acesso, dados, políticas, métricas, padrões e serviços abertos. Por fim, Vicente-Saez e Martinez-Fuentes (2018) fazem uma revisão sistemática para chegar a uma definição comum entre diversos trabalhos e chegam a definir a ciência aberta como *"... um conhecimento transparente e acessível que é compartilhado e desenvolvido através de rede colaborativa"*.

Dados abertos são um dos elementos chave da ciência aberta (Fapesp, n.d.; Unesco, 2021) e, também, encontra várias definições. Para o Open Data Handbook (Open

Data Handbook, n.d.) “dados abertos são dados que podem ser livremente usados, reutilizados e redistribuídos por qualquer pessoa - sujeitos apenas, no máximo, ao requisito de atribuição e compartilhamento”. Apresenta ainda uma definição mais completa e precisa em que enfatiza os aspectos de a) disponibilidade e acesso; b) reuso e redistribuição; e, c) participação universal (qualquer um podendo usar, reusar e distribuir); e conclui indicando a interoperabilidade, isto é, a capacidade de diversos sistemas, organizações e com diferentes tipos de dados trabalharem juntos (interoperar). Nosso foco aqui é o de dados científicos, mas dados abertos também se referem a dados de empresas, governos e associações com poucos setores onde os dados abertos não apresentam algum papel, com destaque para os dados relacionados aos Objetivos de Desenvolvimento Sustentável (ODS) e da Carta de Dados Abertos do G8 (Davies, *et al.* 2019). Especificamente para dados de pesquisa a Unesco traz a seguinte definição (Unesco, 2021): *“dados de pesquisa abertos que incluem, entre outros, dados digitais e analógicos, brutos e processados, e os metadados que os acompanham, bem como pontuações numéricas, registros textuais, imagens e sons, protocolos, código de análise e fluxos de trabalho que podem ser usados abertamente, reutilizados, retidos e redistribuídos por qualquer pessoa, sujeitos a reconhecimento. Os dados de pesquisa abertos estão disponíveis a qualquer momento e de modo amigável, em formato legível às pessoas e por máquina, de acordo com os princípios de boa governança e gestão de dados, notadamente os princípios FAIR (Findable, Accessible, Interoperable e Reusable), apoiados por curadoria e manutenção regulares”*. Este estudo empregará esta definição para orientar o seu direcionamento. A OCDE (OCDE, 2007) ainda recomenda que o acesso dos dados abertos seja preferencialmente baseado na Internet, o que também será levado em conta aqui.

Independente de diferentes definições espera-se da ciência e dos dados abertos uma série de benefícios para a ciência e sociedade como (OCDE, 2016):

1. Acelerar pesquisa científica e torná-la mais eficiente. Por exemplo, reduzindo a duplicação, tornando mais fácil e amplo o acesso.
2. Aumentar a confiabilidade e a qualidade das pesquisas e seus resultados. Por exemplo, possibilitando maior reprodutibilidade e avaliação transparentes.

3. Tornar a ciência mais sensível às necessidades públicas e desafios da sociedade. Por exemplo, tornando-a acessível a pessoas e grupos fora da comunidade científica e com o engajamento público.

Exemplos claros desses benefícios para a ciência e sociedade incluem o projeto do genoma humano e mais recentemente a resposta da ciência à pandemia de COVID-19 acelerando obtenção de soluções para um desafio global (OCDE, 2021).

Disponibilizar dados de forma aberta é uma tradição na área de dados das ciências geo e astrofísicas. O conceito de acesso aberto a dados científicos foi estabelecido já em 1957-1958 (NAP, 2008) como uma forma de minimizar riscos de perda de dados e maximizar a acessibilidade para grande quantidade de dados que eram produzidos por satélites em escala global, e hoje encontram-se inúmeros sites de compartilhamento de dados, incluindo dados e imagens de satélites e de radiotelescópios, com informações do sol e seus fenômenos (solarmonitor.org/), do clima espacial (swpc.noaa.gov/), do sistema global de navegação por satélite (<https://cddis.nasa.gov/archive/gnss/data/daily/>) e da rede de radiotelescópios ALMA (Atacama Large Millimeter Array, em <https://almascience.eso.org/>) para citar apenas alguns.

Grupos de pesquisa também disponibilizam dados abertos de antenas VLF. Destaca-se aqui a iniciativa de Cohen e Golkowski (Cohen e Golkowski, n.d.; Cohen, 2020; Cohen *et. al*, 2019) que iniciaram em 2019 a disponibilização pública de dados VLF, gerados desde os anos 70 até agora, de dezenas de estações gerenciadas pelas universidades de Stanford, Georgia Tech e Colorado-Denver (ver <https://waldo.world/>). O Sudden Ionospheric Disturbances Monitoring Station A118 website (SID, n.d.) é outro exemplo que fornece também dados abertos de VLF.

Nosso projeto tem muito em comum com o projeto WALDO, mas há ao menos três diferenças entre essa iniciativa e o nosso projeto: a principal é de que o nosso projeto não vincula a disponibilização dos dados a um repositório próprio do projeto, o que traz uma redução do custo, uma maior flexibilidade e agilidade para disponibilização dos dados (um dos fatores que levam o projeto WALDO ainda estar, passados 3 anos realizando a transferência de dados); segundo, o projeto WALDO inclui somente

dados gerados por sistemas de antena AWESOME. Nosso projeto inclui dados gerados pelo sistema SAVNET, de diferente formato, e está desenhado para suportar ainda outros formatos de dados; terceiro, a visualização dos dados é desenhada para gerar gráficos apenas sobre demanda. Gráficos gerados são armazenados apenas temporariamente, como um *cache*, caso sejam novamente consultados. Isso permite uma economia de armazenamento (e lembramos que nosso sistema não tem um armazenamento próprio dedicado) comparado ao sistema WALDO, que armazena visualizações de todos os dados, e produzidas previamente, em formato .png (Portable Network Graphics).

2.2. VLF e seu uso no estudo de fenômenos físicos

Este projeto não desenvolve nenhuma pesquisa básica sobre dados de VLF, como seriam um estudo sobre erupções solares ou de detecção tremores sísmicos a partir de sinais de VLF coletados. Entretanto, é importante para o desenvolvimento da plataforma a compreensão de conceitos básicos que envolvem dados de sinais VLF e sua aplicação na pesquisa de certos fenômenos físicos.

2.2.1. Sinal, Propagação, Fontes e Receptoras

Sumariza-se aqui alguns aspectos dos sinais VLF, como eles se propagam, fontes de sinais VLF e antenas receptoras. A maior parte dessas informações encontra-se detalhada em Sizun (2005).

Ondas de rádio frequência encontram-se na faixa de 3kHz a 300GHz. Ondas ou sinais de VLF são ondas de rádio com uma frequência muito baixa (tipicamente 3-30 kHz) comparado a ondas como MF (rádio AM, ~1MHz) e VHF (rádio FM, ~100MHz). Dada a baixa frequência, o comprimento de onda de sinais VLF varia 100-1000Km ($f\lambda = c$, f é a frequência, λ o comprimento de onda e c é a velocidade da luz). Boa parte dos sinais VLF encontram-se dentro de um espectro de ondas audíveis (20Hz-20KHz).

Quanto a propagação ondas de rádio VLF ela ocorre basicamente em três modos (ver Figura 2). O modo direto (*direct wave*) corresponde a parcela transmitida na linha de visão, sem obstáculos entre o transmissor e o receptor. No modo superfície (onda terrestre ou *ground wave*) a onda se propaga interagindo com a superfície condutora da terra e o sinal pode acompanhar a curvatura da terra além do horizonte visível

(~64Km). Finalmente, no modo ionosférico (ou *sky wave*) a onda se propaga sendo refletida pela camada da ionosfera da atmosfera terrestre. Na maior parte dos casos de interesse não há linha de visão entre a antena transmissora e a antena receptora e, portanto, a propagação terrestre e ionosférica são as de maior interesse. Elas permitem que sinais de VLF percorram muito longas distâncias com baixa atenuação tornando-as ideais para transmissões de longa distância. De fato, antenas receptoras capturam sinais VLF gerados em qualquer ponto da terra, podendo uma antena receber sinais de um mesmo transmissor por mais de um sentido da esfera terrestre (por exemplo, os sinais de uma antena sendo recebidos no sentido oeste e leste).

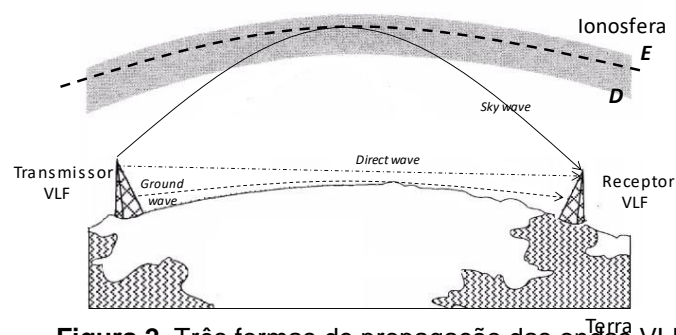


Figura 2. Três formas de propagação das ondas VLF: direta (*direct wave*), terrestre (*ground wave*) e ionosférica (*sky wave*).

A produção de sinais VLF exige antenas verticais de várias centenas de metros (Barr *et al.*, 2000) e grande potência havendo pouco mais de 30 antenas transmissoras no mundo.

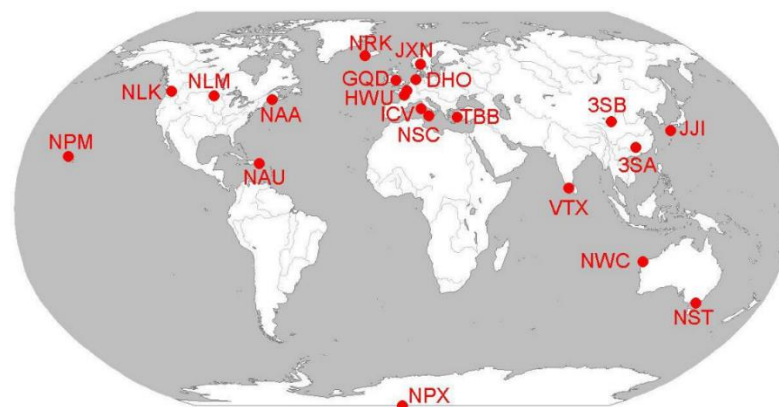


Figura 3. Estações transmissoras VLF conhecidas (Gritsevich *et al.*, 2016).

Ondas VLF são também produzidas na natureza. Raios são uma potente fonte natural de ondas VLF, e antenas de receptoras podem capturar sinais de raios que ocorrem em todo o globo. Ondas VLF também são produzidas por erupções vulcânicas, tornados ou tempestades de areia (Barr *et al.*, 2000).

As estações receptoras, como as estações da rede SAVNET, são constituídas tipicamente de: uma antena receptora com um design específico empregando algumas dezenas ou centenas de metros de cabo receptor, em *loop* ou não (Figura 4); um equipamento receptor de sinais de 15KHz a 80KHz que amplifica e faz a conversão dos sinais analógicos para digital; e um software para aquisição e armazenamento dos dados de amplitude e fase gerados (SID, n.d.).

Uma antena receptora recebe sinais de várias antenas e fontes naturais. O conjunto de todos os sinais constitui os sinais de banda larga (*broadband*). O sinal selecionado de uma transmissora específica é um sinal de banda estreita (*narrowband*), e uma estação receptora seleciona dados *narrowband* de uma ou mais antenas transmissoras. Os dados *narrowband* são também mais empregados.



Figura 4. Antena receptora de VLF da rede SAVNET no Rádio Observatório Pierre Kauffman, em Atibaia/São Paulo. Ela é constituída de dezenas de metros de um fio condutor. Em um design menos eficiente que este esse longo fio pode ser colocado em um *loop* formando uma bobina.

2.2.2. Ionosfera

A atmosfera terrestre pode ser entendida como uma série de camadas concêntricas delimitando diferentes regiões. Camadas até 80-90Km apresentam nitrogênio e oxigênio em proporções constantes, enquanto os gases leves, como nitrogênio, hidrogênio e hélio são predominantes em alturas superiores. A ionosfera é uma camada de elétrons e íons eletricamente carregados (átomos e moléculas) que se estende da altura de cerca de 60km até 1000km, e sua natureza carregada se deve a produção de plasma (elétrons, íons) principalmente devido à radiação (ultravioleta) solar (PhysicsOpenLab, 2020). Essas cargas presentes em quantidades grandes o suficiente para refletir as ondas de rádio (Sizun, 2005).

A ionosfera também pode ser dividida em camadas (D,E,F). A camada D surge na região de 70-90km de altura e é a camada ionizada (carregada) da ionosfera. Essa é uma camada dinâmica produzida durante o dia com a ionização pelos raios solares e que vai progressivamente desaparecendo à noite, à medida que não incidem mais raios solares sobre a ionosfera. Desse modo, durante o dia, a reflexão da propagação ionosférica das ondas VLF (*sky wave*) se dá na camada D, mais abaixo, enquanto à noite a reflexão se dá na parte inferior da camada E (Figura 5. (a) e (b)).

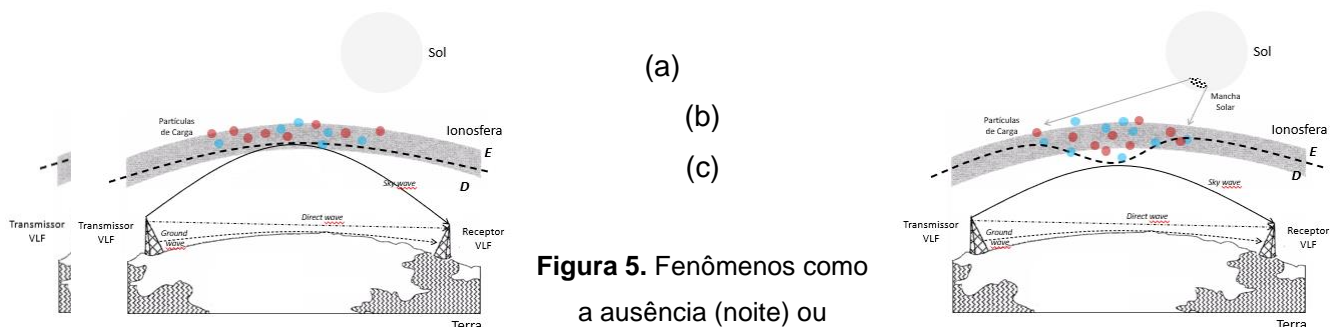


Figura 5. Fenômenos como a ausência (noite) ou

presença (dia) da luz do dia, e explosões solares que alteram a composição da ionosfera. Isso causa interferências na propagação ionosférica das ondas VLF produzindo uma 'assinatura' desses fenômenos na amplitude e fase dos sinais que pode, então, ser empregada para identificar e analisar esses fenômenos.

Quaisquer outros fenômenos que, a exemplo da luz do dia, afetem a composição da ionosfera causam interferência na propagação das ondas VLF e esta *assinatura* pode então ser empregada como uma medida indireta para a análise do fenômeno. Essas

perturbações da ionosfera podem ter origem em eventos terrestres ou extra-terrestres, e serem periódicas ou efêmeras. Entre os eventos extra-terrestres, o sol é o principal responsável por alterações na ionosfera e perturbações periódicas são causadas pela rotação da terra (com a presença ou não da luz do dia), revolução da terra (que alteram a incidência dos raios solares) e o ciclo solar (de maior atividade do sol a cada 11 anos). Dentre as perturbações efêmeras (ou repentinas) as mais importantes são as causadas por explosões solares (Figura 5 (c)). Essas explosões produzem uma grande quantidade de raios-x com um forte efeito de ionização da região D impactando os sistemas de comunicação de rádio e satélite. Dentre os fenômenos terrestres que afetam a composição da ionosfera encontramos raios, erupções vulcânicas e eventos sísmicos e fenômenos atmosféricos como tempestades e ciclones (Nina, 2000). Quaisquer desses fenômenos podem ser estudados do mesmo modo a partir da assinatura que essas perturbações causam nos sinais de VLF medidos.

3. METODOLOGIA DA PESQUISA

Esta pesquisa é de natureza aplicada, com a abordagem quantitativa, finalidade propositiva e com pesquisas documentais e bibliográficas que dever proporcionar, ao final, uma plataforma aberta (acesso público) e amplamente acessível (web) para os dados das antenas VLF da rede SAVNET do CRAAM, idealizada para que possa ser adaptável para outros tipos de dados e permitir a fácil inclusão de quaisquer outros repositórios de dados VLF semelhantes.

Uma fase preliminar deste projeto, já quase concluída, contempla o entendimento dos dados, a avaliação de plataformas semelhantes já desenvolvidas, a seleção de ferramentas a serem empregadas e um desenho da solução.

A plataforma a ser desenvolvida deve suportar o compartilhamento de dados até 4TB, suportando inicialmente dados de VLF produzidos por dois sistemas diferentes, AWESOME (.mat) e SAVNET (.fits), que compõe a rede de antenas receptoras de VLF do CRAAM. No seu desenvolvimento deve empregar predominantemente Javascript (front-end) e Python (back-end), buscando-se empregar as melhores práticas de desenvolvimento e fazendo uso de frameworks e bibliotecas (como Vue, Pandas, Matplotlib etc.) que facilitem a criação da plataforma. A plataforma será

construída a fim de suportar diferentes fontes de dados e deve, inicialmente, suportar ao menos armazenamento local (file system) e um armazenamento em nuvem (Google Drive, AWS S3, owncloud ou similar). Uma base MongoDB servirá para o armazenamento dos metadados que mapeiam os diferentes repositórios de dados. A Figura 6 ilustra um desenho esquemático da solução.

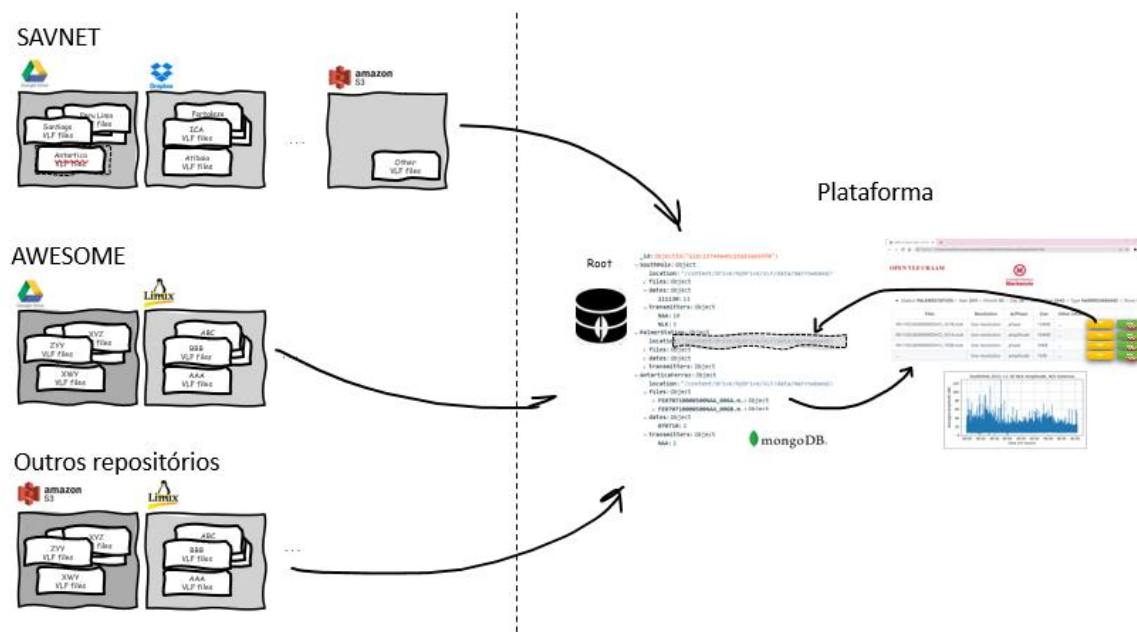


Figura 3. Esquemático da plataforma proposta para o acesso aberto aos dados. Ela permite o acesso aos dados VLF, de sistemas SAVNET e AWESOME, armazenados em diferentes repositórios e tecnologias. Os metadados que mapeiam as informações de cada repositório encontram-se em um banco de dados (MongoDB) da solução. A interface (Web) fornece ferramentas de busca, seleção, visualização e acesso aos dados.

As demais etapas do projeto, além do desenvolvimento e teste da plataforma, consistem em: a) Revisão da literatura sobre ciência aberta e dados abertos, e soluções atuais, bem como de uma literatura sobre VLF e as pesquisas sobre esses dados para suporte ao conteúdo informativo da plataforma; b) Criação do conteúdo informativo da plataforma (informações sobre os dados, tutoriais de uso, páginas com informações sobre VLF e aplicações, referências de trabalhos relacionados e informações de divulgação científica); c) Movimentação e organização dos dados,

hoje em HDs e DVDs para um ou mais dispositivos de armazenamento com acesso via a plataforma; d) Publicação da aplicação na Web; d) Redação e publicação do trabalho final e resultados.

O cronograma de cada uma dessas etapas é apresentado a seguir.

4. CRONOGRAMA

As atividades desta pesquisa estão organizadas da seguinte forma:

- a. **Fase Preliminar.** Contempla o entendimento dos dados, avaliação de plataformas semelhantes já desenvolvidas, a seleção de ferramentas a serem empregadas e um desenho inicial da solução.
- b. **Revisão bibliográfica.** Revisão da literatura sobre ciência aberta e dados abertos, e soluções atuais; Revisão da literatura da literatura sobre VLF e as pesquisas que esses dados envolvem para o conteúdo informativo do site.
- c. **Desenvolvimento do Conteúdo Informativo.** Desenvolvimento do conteúdo informativo do site incluindo informações para usuários dos dados e de divulgação científica.
- d. **Desenvolvimento e Testes da Plataforma.** Criação de um protótipo e desenvolvimento da aplicação de forma modular e incremental; inclui os códigos de front-end e back-end; popular base de dados; interações com os participantes do projeto para refinar a solução; etc.
- e. **Organização dos dados.** Movimentação e Organização dos dados atuais (HDs e DVDs) para um ou mais dispositivos de armazenamento com acesso pela plataforma.
- f. **Publicação da Aplicação na Web.** Definição da plataforma final e distribuição da aplicação.
- g. **Publicação Final do Projeto.** Redação e publicação do trabalho final e resultados, indicando possíveis caminhos para continuidade do projeto.

Com o seguinte cronograma:

Atividades	2021	2022											
	-	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
a. Fase Preliminar. Contempla o entendimento dos dados, avaliação de plataformas semelhantes já desenvolvidas, a seleção de ferramentas a serem empregadas e um desenho inicial da solução.													
b. Revisão bibliográfica. Revisão da literatura sobre ciência aberta e dados abertos, e soluções atuais; Revisão da literatura da literatura sobre VLF e as pesquisas que esses dados envolvem para o conteúdo informativo do site.													
c. Desenvolvimento do Conteúdo Informativo. Desenvolvimento do conteúdo informativo do site incluindo informações para usuários dos dados e de divulgação científica.													
d. Desenvolvimento e Testes da Plataforma. Criação de um protótipo e desenvolvimento da aplicação de forma modular e incremental; inclui os códigos de front-end e back-end; popular base de dados; interações com os participantes do projeto para refinar a solução; etc.													
e. Organização dos dados. Movimentação e Organização dos dados atuais (HDs e DVDs) para um ou mais dispositivos de armazenamento com acesso pela plataforma.													
f. Publicação da Aplicação na Web. Definição da plataforma final e distribuição da aplicação.													
g. Publicação Final do Projeto. Redação e publicação do trabalho final e resultados, indicando possíveis caminhos para continuidade do projeto.													

REFERÊNCIAS

UNESCO. UNESCO Recommendation on Open Science (2021). SC-PCB-SPP/2021/OS/UROS.

<https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=em>.

FAPESP. Gestão de Dados (2017) <https://fapesp.br/gestaodedados>

FAPESP. Open Science (n.d.) <https://www.fapesp.br/openscience/>

OECD. Making Open Science a Reality, (2015) OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris, <https://doi.org/10.1787/5jrs2f963zs1-en>

OECD. OECD Principles and Guidelines for Access to Research Data from Public Funding (2007). <https://www.oecd.org/sti/inno/38500813.pdf>

EC. European Commission, Directorate-General for Research and Innovation, Hessels, L., Koenig, L., Diederer, P., Perspectives on the future of open science : effects of global variation in open science practices on the European research system, Publications Office of the European Union, (2021),

<https://data.europa.eu/doi/10.2777/054281> (EN-PDF) ISBN 978-92-76-41133-8 doi: 10.2777/054281

Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, 88, 428–436. doi:10.1016/j.jbusres.2017.12.043

Foster. What is Open Science? Introduction. (n.d.).
<https://www.fosteropenscience.eu/content/what-open-science-introduction>

Open Data Handbook. What is Open Data? (n.d.) <https://opendatahandbook.org/>

Davies, T., Walker, S. B., Rubinstein, M., & Perini, F. The state of open data: Histories and horizons. (2019). *African Minds*. <https://stateofopendata.od4d.net/>

OCDE. Open Data in action Initiatives during the initial stage of the COVID-19 pandemic (2021). <https://www.oecd.org/gov/digital-government/open-data-in-action-initiatives-during-the-initial-stage-of-the-covid-19-pandemic.pdf>

NAP, Committee on Scientific Accomplishments of Earth Observations from Space, National Research Council (2008). *Earth Observations from Space: The First 50 Years of Scientific Achievements*. The National Academies Press. p. 6. doi:10.17226/11991. ISBN 978-0-309-11095-2. Acesso: <https://nap.nationalacademies.org/read/11991/chapter/2>

Cohen, M.; Golkowski, M. WALDO database (n.d.) <https://waldo.world>.

Cohen, M. Returning Lightning Data to the Cloud (2020) <https://eos.org/science-updates/returning-lightning-data-to-the-cloud>

Cohen, M., Golkowski, M., Inan, U., & DeSilva, J. (2019). A Massive Public Repository of Global ELF/VLF Radio Data. In *AGU Fall Meeting Abstracts* (Vol. 2019, pp. SA33B-3156).

SID. Sudden Ionospheric Disturbances Monitoring Station A118 (n.d.).
<https://sidstation.loudet.org/home-en.xhtml>

Nina, A. (2020). Application of Databases Collected in Ionospheric Observations by VLF/LF Radio Signals. In Knowledge Discovery in Big Data from Astronomy and Earth Observation (pp. 419-434, chapter 23). Elsevier. <https://doi.org/10.1016/B978-0-12-819154-5.00035-7>

Raulin, J. P., Correia de Matos David, P., Hadano, R., Saraiva, A. C., Correia, E., & Kaufmann, P. (2009). The South America VLF NETwork (SAVNET). *Earth, Moon, and Planets*, 104(1), 247-261.

Raulin, J. P., Bertoni, F. C., Gavilán, H. R., Guevara-Day, W., Rodriguez, R., Fernandez, G., ... & Hadano, R. (2010). Solar flare detection sensitivity using the South America VLF Network (SAVNET). *Journal of Geophysical Research: Space Physics*, 115(A7).

Correia, E. (2011). Study of Antarctic-South America connectivity from ionospheric radio soundings. *Oecologia Australis*, 15(1), 32-39.

Bertoni, F. C. P., Raulin, J. P., Gavilán, H. R., Kaufmann, P., Rodriguez, R., Clilverd, M., ... & Fernandez, G. (2013). Lower ionosphere monitoring by the South America VLF Network (SAVNET): C region occurrence and atmospheric temperature variability. *Journal of Geophysical Research: Space Physics*, 118(10), 6686-6693.

Tacza, J., Raulin, J. P., Macotela, E., Norabuena, E., Fernandez, G., Correia, E., ... & Harrison, R. G. (2014). A new South American network to study the atmospheric electric field and its variations related to geophysical phenomena. *Journal of Atmospheric and Solar-Terrestrial Physics*, 120, 70-79.

Gritsevich, M., Vinkovic, D., Schwarz, G., Nina, A., Koschny, D., & Lyytinen, E. (2016, December). Meteor Observations as Big Data Citizen Science. In AGU Fall Meeting Abstracts (Vol. 2016, pp. ED21C-0794).

Sizun, H. (2005). Radio wave propagation for telecommunication applications. Berlin/Heidelberg/New York: Springer.

Barr, R., Jones, D. L., & Rodger, C. J. (2000). ELF and VLF radio waves. Journal of Atmospheric and Solar-Terrestrial Physics, 62(17-18), 1689-1718.

PhysicsOpenLab. VLF Receiver for SID Monitoring (2020).

<https://physicsopenlab.org/2020/05/07/vlf-receiver-for-sid-monitoring/>