

Sessão Lightning Talks

29/08/2018



- 1. RDP - Rede de Dados de Pesquisa Brasileira**
Rafael da Rocha (UFRGS) - 5 min.
- 2. EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications**
Walter dos Santos Filho (UFMG) - 5 min.
- 3. Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing**
Francisco Brasileiro (UFCG) - 5 min.
- 4. GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede**
Marinho Barcellos (UFRGS) - 5 min.
- 5. GT-AMPTo – Autenticação Multi-fator para Todos**
Michelle Silva Wangham (UNIVALI) - 5 min.



1. O que é o **produto em desenvolvimento**?
2. Quem são os **usuários alvo**?
3. Qual é o principal **resultado** alcançado na **fase atual**?
4. Como estes resultados podem contribuir para a construção de **Patrimônio digital, de dados abertos ou de tecnologias de memória**?
5. Qual a sua **visão de futuro** para a aplicação dos resultados no curto prazo (6 a 18 meses)?



- 1. RDP - Rede de Dados de Pesquisa Brasileira**
Rafael da Rocha (UFRGS) - 5 min.
- 2. EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications**
Walter dos Santos Filho (UFMG) - 5 min.
- 3. Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing**
Francisco Brasileiro (UFCG) - 5 min.
- 4. GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede**
Marinho Barcellos (UFRGS) - 5 min.
- 5. GT-AMPTo – Autenticação Multi-fator para Todos**
Michelle Silva Wangham (UNIVALI) - 5 min.

Rede de Dados de Pesquisa Brasileira

Rafael Port da Rocha

Universidade Federal do
Rio Grande do Sul

O que é o produto em desenvolvimento?

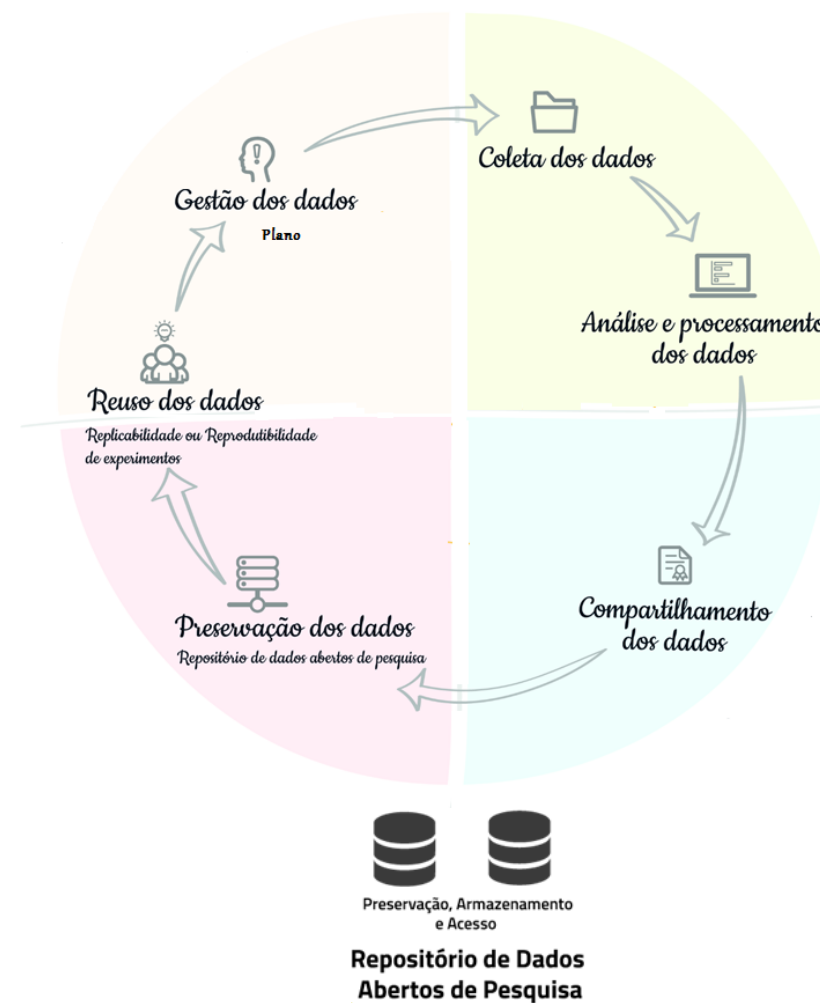


Protótipo de Repositório de Dados da Pesquisa

Visa:

Promover a Preservação e o Compartilhamento do Dado de Pesquisa no contexto do seu Ciclo de Vida

Subsidiar ações de Acesso Aberto a Dados de Pesquisa (AADP) no Brasil



Quem são os usuários alvo?

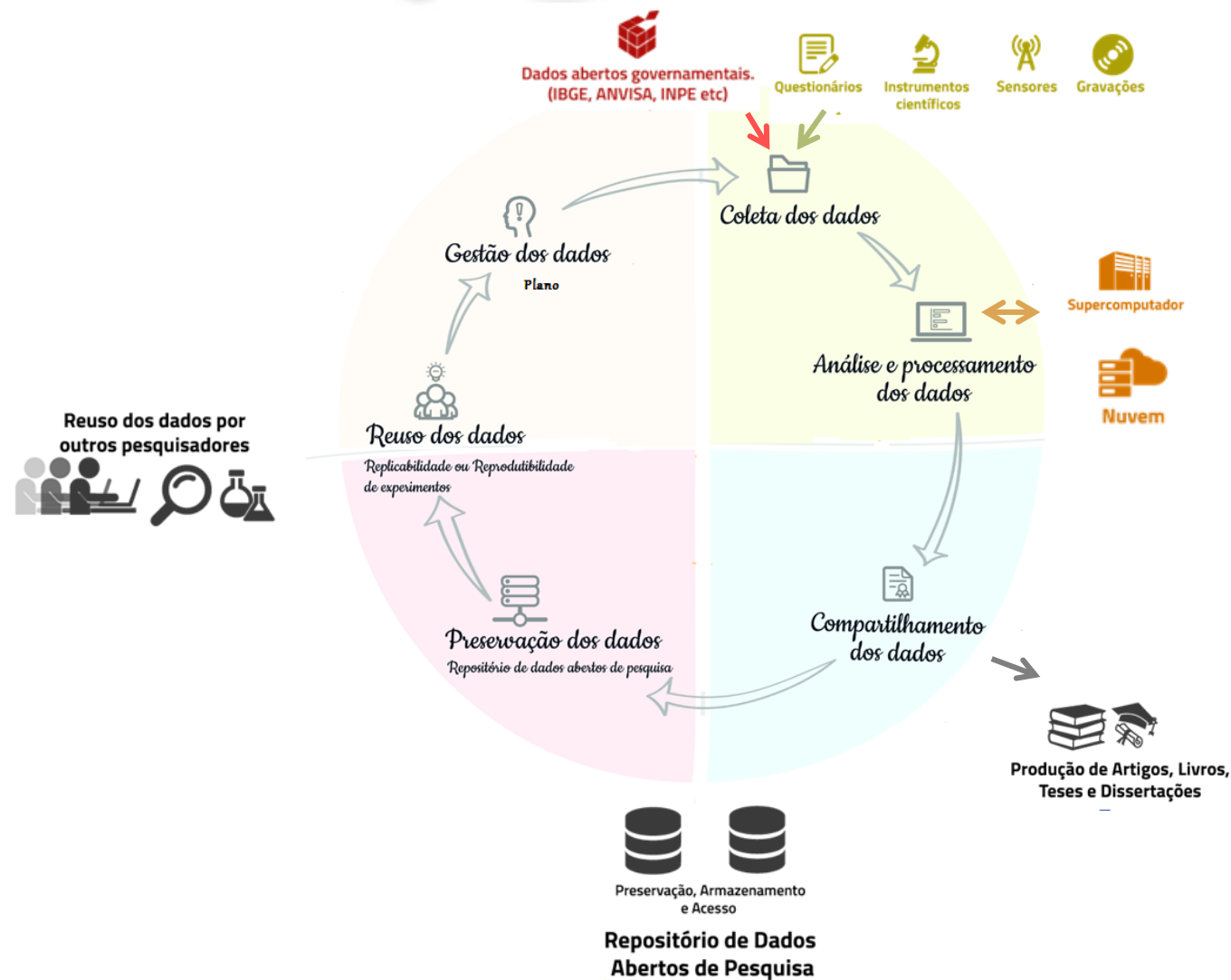
Usuários

Pesquisadores do Brasil,

que produzem e consomem
Dados da Pesquisa:

*Planejam, Coletam, Analisam,
Processam, Compartilham e
Reusam Dados*

Comunidade que promove AADP no
Brasil

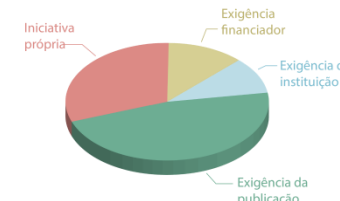


Qual é o principal resultado alcançado na fase atual?

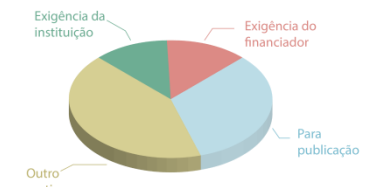


- **Práticas e Percepções sobre AADP no Brasil**
 - Questionário enviado a pesquisadores do Brasil (PPGs, Grupos de Pesquisa)
 - 4.735 respostas
- **Soluções Tecnológicas para Compartilhamento de Dados**
 - Ambiente OAIS, Repositório Confiável
 - Princípios FAIR, Identificação, Proveniência, Metadados, Citação, Representação, Versões
 - Dataverse, Dspace, CKAN, EUDAT
- **Desenvolvimento de Protótipo de Repositório**
 - Repositório para a Área de Redes de Computadores
- **Portal web para a comunidade AADP**
 - <https://dadosdepesquisa.rnp.br>

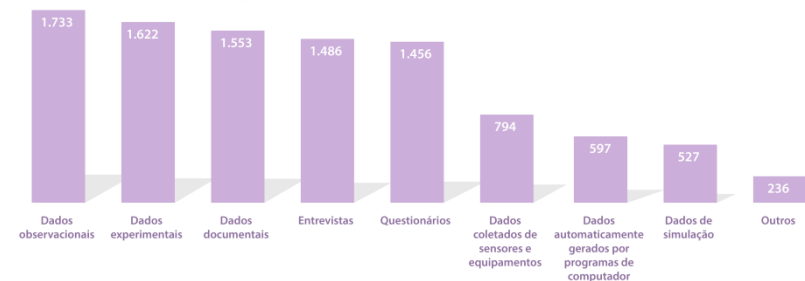
MOTIVOS PARA COMPARTILHAMENTO



MOTIVOS PARA ELABORAÇÃO DE PLANO DE GESTÃO DE DADOS



TIPOS DE DADOS DE PESQUISA



<https://dadosdepesquisa.rnp.br/>

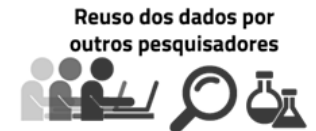
Findable
Accessible
Interoperable
Reusable



- Preservar ativos digitais e a memória da pesquisa brasileira
- Promover o compartilhamento de dados da pesquisa e colher os seus benefícios¹:
 - Permitir a reprodução ou a verificação da pesquisa
 - Disponibilizar resultados de pesquisas financiadas com fundos públicos
 - Possibilitar que outros façam novos questionamentos sobre os dados
 - Avançar no Estado Arte da pesquisa e inovação
- Subsidiar ações ou programas nacionais para AADP



Repositório de Dados
Abertos de Pesquisa



Qual a visão de futuro para aplicação dos resultados no curto prazo (6 a 18 meses)?

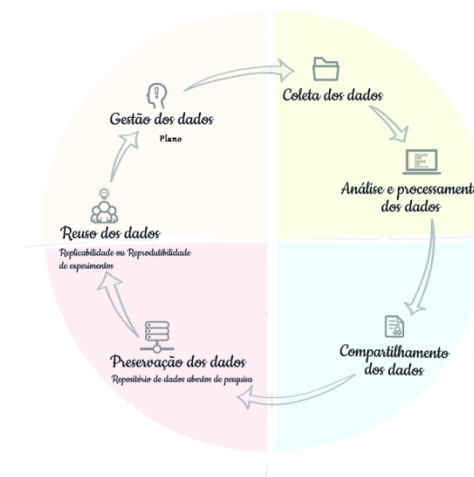


Hoje/Logo:

- Oferta de um Protótipo de Repositório de Dados de Pesquisa para a Área de Redes de Computadores.
- Domínio de Soluções Tecnológicas
- Conhecimento de Práticas e Percepções de Usuários
- Portal AADP

Amanhã:

- Repositório atendendo a uma grande diversidade de dados, provenientes de várias áreas (dados da cauda longa)
- Ações e programas para promover AADP
- Preservação do Patrimônio da Pesquisa
- Memória dos Dados de Pesquisa e seus Ciclos de Vida
- Comunidade AADP



Obrigado

Rafael Port da Rocha

Rafael.rocha@ufrgs.br



1. **RDP - Rede de Dados de Pesquisa Brasileira**
Rafael da Rocha (UFRGS) - 5 min.
2. **EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications**
Walter dos Santos Filho (UFMG) - 5 min.
3. **Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing**
Francisco Brasileiro (UFCG) - 5 min.
4. **GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede**
Marinho Barcellos (UFRGS) - 5 min.
5. **GT-AMPTo – Autenticação Multi-fator para Todos**
Michelle Silva Wangham (UNIVALI) - 5 min.

Lemonade: Computação de alto desempenho para processamento de base de dados abertos (CAPEs)



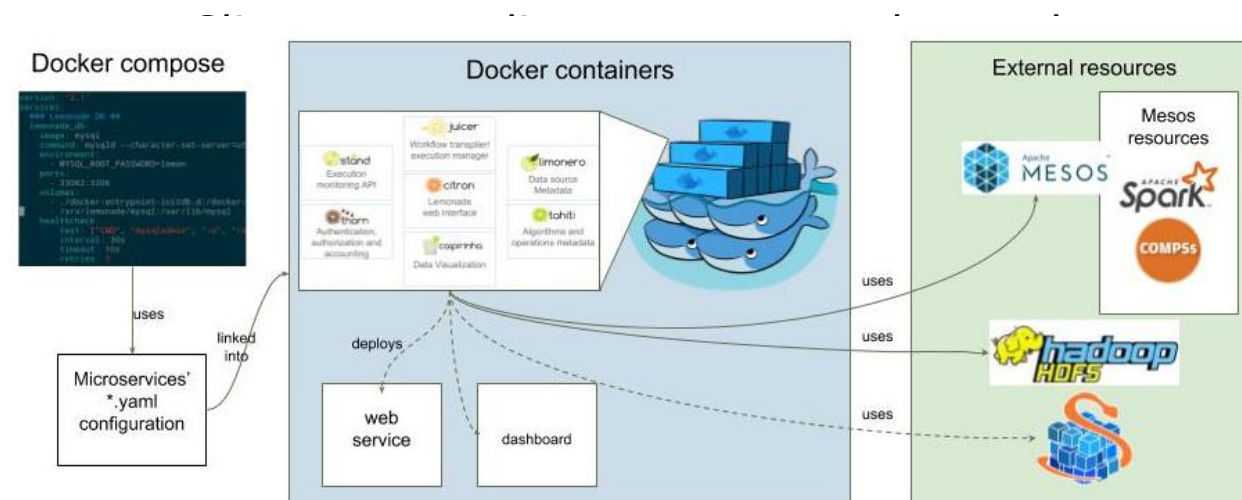
Walter dos Santos Filho

UFMG / EUBra-BIGSEA



Lemonade - Live Exploration and Mining Of a Non-trivial Amount of Data from Everywhere

- Lemonade é uma plataforma para a criação e execução de fluxos de processamento e mineração de dados.
- Disponibiliza funções de processamento e algoritmos de aprendizado de máquina sob a forma de tarefas em um fluxo (workflow).
- Premissa: especialistas no negócio podem obter resultados a partir dos dados, sem necessariamente ter que programar ou manter a infraestrutura

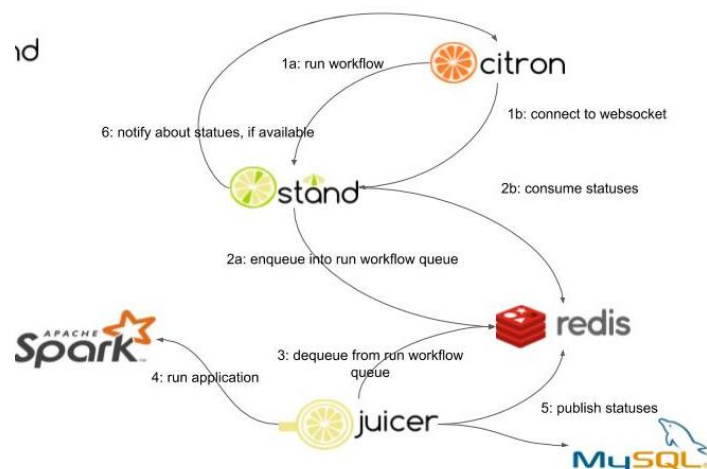




Usuários alvo e porque usar

- Estudantes;
- Cientistas de dados;
- Especialistas em diversas áreas do conhecimento que desejam uma ferramenta para processamento de dados;
- Empresas que desejam iniciar a exploração dos seus dados.

- O Lemonade se apoia em tecnologias consagradas atualmente, tais como Apache Spark, Scikit Learn, Matplotlib, Keras e Tensorflow e várias outras;
- Ao usar fluxos, permite ao usuário a construção de uma solução em um nível mais alto de abstração;
- Integrado com ferramentas de visualização e exploração de dados;
- Permite exportar o código e usá-lo fora da plataforma;
- Código aberto;



Qual é o principal resultado alcançado na fase atual?



Resultados

Acadêmicos:

- 4 artigos científicos;
- 1 dissertação de mestrado;
- Pelo 2o. ano, usado em disciplinas na UFMG (mineração de dados e introdução à ciência dos dados) e na ENAP (Escola Nacional de Administração Pública) como ferramenta de ensino.

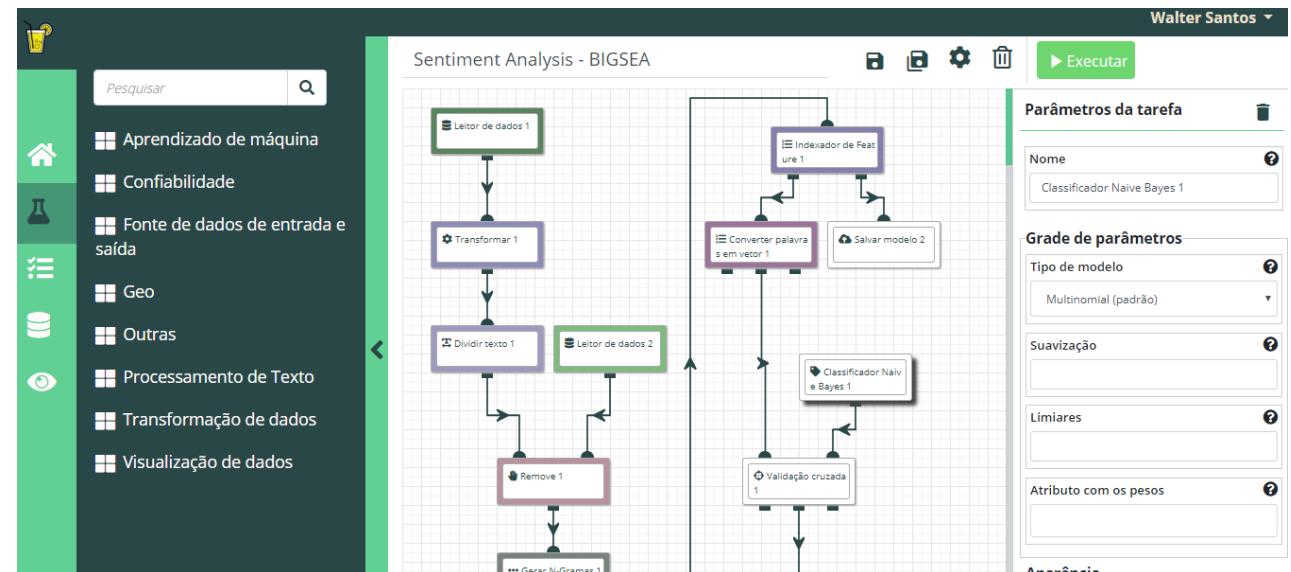
Negócios

- Início de parcerias com Ministério da Saúde e RNP;
- Interesse por parte do Tribunal de Contas da União e empresas privadas de Belo Horizonte.

Como estes resultados podem contribuir para a construção de patrimônio digital, de dados abertos ou de tecnologias de memória

Lemonade pode contribuir:

- Como ferramenta de processamento de dados massivos;
- Como ferramenta de ensino;
- Como uma ferramenta que suporte boas práticas atuais relacionados à ciência dos dados (reprodutibilidade, justiça algorítmica, transparência e interpretabilidade de resultados, privacidade) ou aderentes a leis (GDPR, LGPD);
- Como um repositório de bases de dados e fluxos de trabalho.



Qual a visão de futuro para aplicação dos resultados no curto prazo (6 a 18 meses)?



- O Lemonade está sendo evoluído no projeto Atmosphere (encerra-se em nov/2019):
- Integração com bibliotecas de aprendizado profundo (deep learning) e execução em GPUs;
 - Integração com tecnologias de nuvem para escalabilidade;
 - Disponibilização dos fluxos como serviços web (permitir que um fluxo seja usado como uma API);
 - Para o Atmosphere: apoiar o desenvolvimento de uma aplicação para análise de eco-cardiogramas com a finalidade de diagnosticar a febre reumática (doença que pode afetar as articulações, a pele e até mesmo órgãos vitais, como o coração e o cérebro).
 - Parceria com a RNP para processamento dos dados da CAPES (em andamento)
 - Parceria com o Ministério da Saúde para análise dos dados do Datasus (em andamento)

Obrigado

Walter dos Santos Filho



walter@dcc.ufmg.br



1. **RDP - Rede de Dados de Pesquisa Brasileira**
Rafael da Rocha (UFRGS) - 5 min.
2. **EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications**
Walter dos Santos Filho (UFMG) - 5 min.
3. **Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing**
Francisco Brasileiro (UFCG) - 5 min.
4. **GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede**
Marinho Barcellos (UFRGS) - 5 min.
5. **GT-AMPTo – Autenticação Multi-fator para Todos**
Michelle Silva Wangham (UNIVALI) - 5 min.

Projeto ATMOSPHERE: objetivos e resultados parciais

Francisco Brasileiro

Universidade Federal de Campina Grande



Objetivos do projeto ATMOSPHERE:

- O projeto ATMOSPHERE visa o projeto e o desenvolvimento de uma **plataforma** para implementar **serviços confiáveis** (*trustworthy*), implantados em um ambiente de **provedores de computação na nuvem federados**
 - Considerando um **amplo espectro** de propriedades e medidas de confiabilidade:
 - Segurança (*security*), Garantia de privacidade (*privacy*), Coerência (*coherence*), Isolamento (*isolation*), Estabilidade (*stability*), Equidade (*fairness*), Transparência (*transparency*) e a Confiança no funcionamento (*dependability*)
- O projeto ATMOSPHERE fornecerá:
 - Um conjunto de métricas, regras, testes e procedimentos para **avaliar automaticamente as várias dimensões de confiabilidade** de uma aplicação
 - Uma **plataforma para monitorar e medir essas métricas**, tanto durante o projeto, como durante a execução das aplicações
 - Serviços para o **armazenamento e o processamento confiáveis de dados**.
 - **Implantação automática** e configuração de **infraestruturas complexas** a partir de receitas de configuração de software.



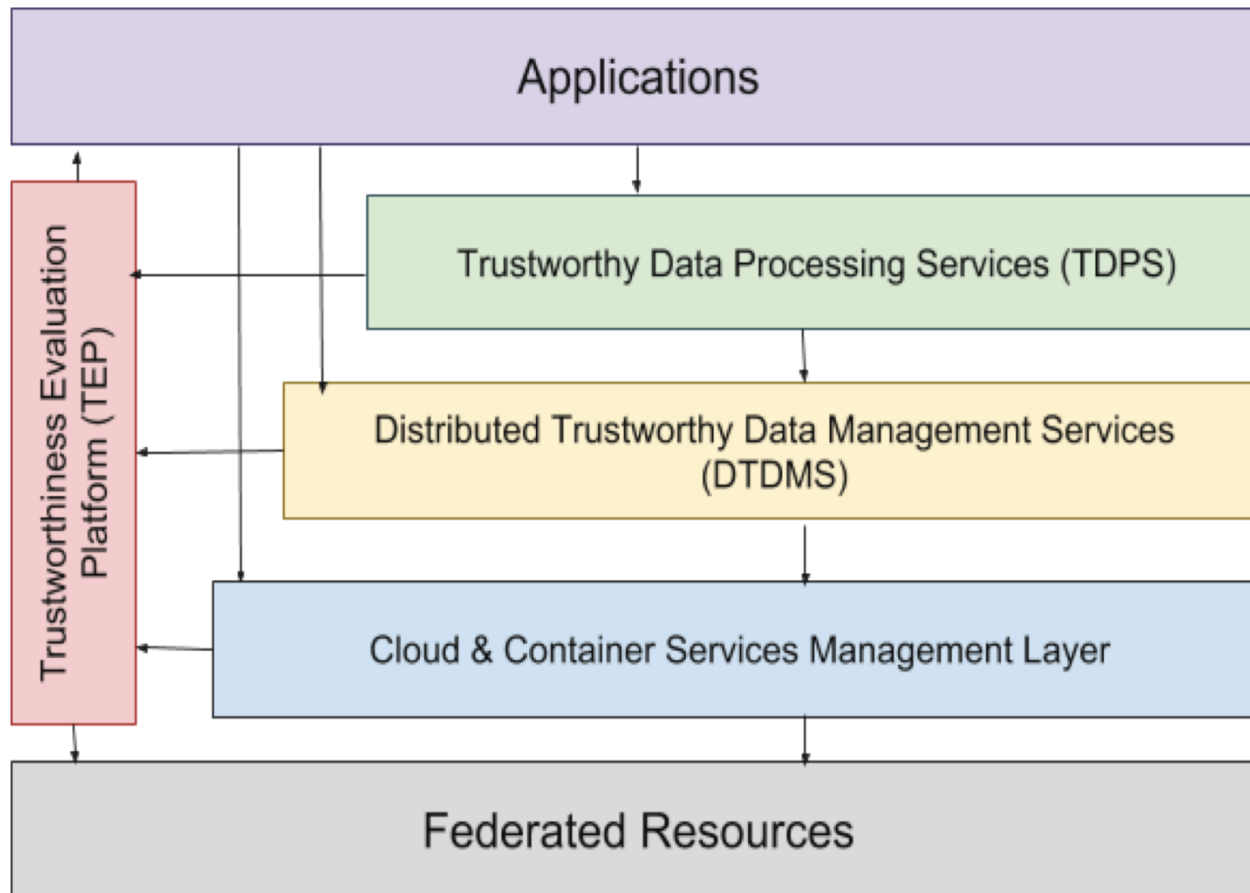
Os resultados do projeto afetarão diferentes “*stakeholders*”

- Provedores de computação na nuvem
 - Poderão utilizar as tecnologias e produtos desenvolvidos pelo projeto para **oferecer serviços de armazenamento e processamento de dados de forma mais confiável**
 - Poderão **medir a confiabilidade dos serviços** oferecidos
- Startups e PMEs
 - As tecnologias desenvolvidas poderão abrir **oportunidades de inovação** para essas empresas
- Provedores de serviços hospedados em infraestruturas de computação na nuvem
 - Poderão **ofertar serviços que atualmente não podem ser ofertados** pelas limitações de confiabilidade dos serviços oferecidos pelos provedores de computação na nuvem
- Usuários de aplicações relacionadas a áreas que normalmente tratam com dados sensíveis
 - Ex. área de **saúde, segurança pública**, etc.

Qual é o principal resultado alcançado na fase atual?



Resultados parciais



- Medidas de confiabilidade
 - Foram definidas métricas para avaliar diferentes dimensões de confiabilidade
- Processamento confiável de dados
 - Melhorias na ferramenta **Lemonade** para processamento de grandes massas de dados, considerando novas dimensões de confiabilidade
- Armazenamento confiável de dados
 - Utilização de hardware especializado para aumentar a segurança e a privacidade dos dados processados
- Infraestrutura
 - Ampliação dos recursos do middleware **Fogbow** para federação de provedores de nuvem

Contribuições para o processamento de dados abertos

- Aumentar a confiabilidade dessas aplicações, sobretudo aquelas que são implantadas em ambientes distribuídos, possivelmente distribuídos em regiões geográficas sujeitas a diferentes ordenamentos jurídicos
- Ampliar os tipos de dados que poderão ser processados, ao permitir o processamento de dados sensíveis, respeitando os requisitos de segurança e privacidade associados aos mesmos
- Tratar de dimensões de confiabilidade fundamentais para aplicações de processamento massivo de dados (*Big Data*), até então pouco exploradas na literatura científica:
 - Ex. transparência e equidade dos modelos de classificação

Qual a visão de futuro para aplicação dos resultados no curto prazo (6 a 18 meses)?



Próximos passos

- O projeto ATMOSPHERE tem vigência até outubro de 2019
- Até outubro de 2018 deveremos ter uma primeira versão da plataforma demonstrável
- Até outubro de 2019 teremos uma versão estável da plataforma dando suporte a uma aplicação de saúde que será utilizada como prova de conceito

Obrigado!

Francisco Vilar Brasileiro

 fubica@computacao.ufcg.edu.br



1. RDP - Rede de Dados de Pesquisa Brasileira
Rafael da Rocha (UFRGS) - 5 min.
2. EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications
Walter dos Santos Filho (UFMG) - 5 min.
3. Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing
Francisco Brasileiro (UFCG) - 5 min.
4. **GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede**
Marinho Barcellos (UFRGS) - 5 min.
5. GT-AMPTo – Autenticação Multi-fator para Todos
Michelle Silva Wangham (UNIVALI) - 5 min.

GT-IPÊ Analytics

Marinho Barcellos

UFRGS



Sistema de análise do estado da rede baseado na transformação de dados brutos de monitoramento em informações de mais alto nível

Facilita gerenciamento e a operação das redes, tanto no nível mais baixo (operação dos dispositivos), quanto no mais alto (identificação/priorização de quais pontos podem ser melhorados)

Amplia visão das equipes de gerência da rede para potenciais problemas que necessitam de correção (ativo/reactivo)

Interação simples e intuitiva, acessível por meio de qualquer navegador moderno

Sistema dividido em front-end e back-end

Front-end é interface com “usuário”

Backend processa conjuntos de dados

Foco original em conjuntos de dados pré-existent, minimizando sobrecarga à RNP

Ao invés de um *produto*, poderá ser vendido como um **serviço**, cujo uso será por assinatura



Operadores da rede da RNP e seus clientes, comunidade científica

Equipes de Gerência da rede da RNP

Provedores de rede: Provedores de Serviço da Internet (ISPs) ou Redes Nacionais de Ensino e Pesquisa (NRENS)

Redes corporativas ou acadêmicas (empresas de médio e grande porte, universidades)

Comunidade científica para gerar e processar conjuntos de dados

Os clientes do **serviço**, mediante assinatura periódica

Qual é o principal resultado alcançado na fase atual?



Sistema funcional com interface gráfica para inferências de rede

Adição de perfis diferentes, de acordo com os papéis diferentes de usuários de uma mesma entidade cliente (operadores, gerentes e administradores de um ISP ou NREN, por exemplo)

Melhorias visuais e ampliação das inferências

Melhorias no desempenho (tempo de resposta a consultas) e na disponibilidade do sistema

Refinamentos necessários à criação de um produto (instalação e configuração automatizada, criação de VMs para lançamento de “appliance” ou contratação de serviço na nuvem

Melhorias nos processos de coleta de dados da rede

Cooperação com o CT-Mon para disponibilização de dados à comunidade científica

Geração de conjuntos de dados disponibilizados à comunidade científica

Sistema foi desenvolvido, desde sua concepção, para preservar todos os dados coletados e fazer análises históricas destes

Assim, ajuda a preservar o estado da rede e sua evolução ao longo do tempo

Futuramente, pode contribuir com as iniciativas como a do RDP e as discussões do CT-Mon, no sentido de fornecer a base necessária para disponibilização de dados à comunidade científica

Há vários desafios para fazer isso de maneira segura e eficiente, como coletar, organizar, consolidar, sanitizar e disponibilizar consultas de forma correta

Qual a visão de futuro para aplicação dos resultados no curto prazo (6 a 18 meses)?



Curto prazo: fornecer uma nova ferramenta de trabalho para engenharia e operações, através da melhora na identificação e correção de problemas e na previsão de comportamentos

Médio prazo: criar um serviço essencial para a operação de redes de larga escala (ISPs, NRENS) e para redes corporativas e acadêmicas

Gratidão

Marinho Barcellos (UFRGS)

marinho@inf.ufrgs.br



1. **RDP - Rede de Dados de Pesquisa Brasileira**
Rafael da Rocha (UFRGS) - 5 min.
2. **EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications**
Walter dos Santos Filho (UFMG) - 5 min.
3. **Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing**
Francisco Brasileiro (UFCG) - 5 min.
4. **GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede**
Marinho Barcellos (UFRGS) - 5 min.
5. **GT-AMPTo – Autenticação Multi-fator para Todos**
Michelle Silva Wangham (UNIVALI) - 5 min.

GT-AMPTo: Grupo de Trabalho de Autenticação Multi-fator para Todos

Michelle Wangham

UNIVALI



Solução que permite as instituições usuárias da RNP realizarem autenticação de seus usuários com mais de um fator de autenticação

- **O que você sabe:** usuário/senha, PIN
- **O que você possui:** dispositivo móvel, *token* criptográfico
- **O que você é:** impressão digital, íris, face e voz

- **CAFe (hoje):** usuário/senha
- **CAFe + solução GT-AMPTo** = opções de segundo fator
 - Módulo para Provedores de Identidade Shibboleth
 - Gerencia o ciclo de vida do segundo fator

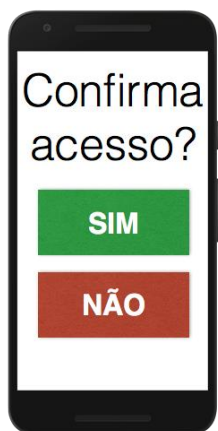


- Instituições que possuem provedor de identidade *Shibboleth*
 - Todas instituições usuárias da RNP que estão na federação CAFe
- Usuários de instituições de ensino e pesquisa
 - Evitar problemas com ataques de roubo de senhas (*phishing* ou *malwares*)

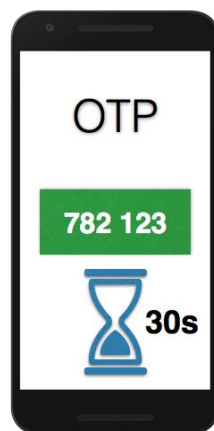
Qual o principal resultado alcançado na fase atual?



Primeiro fator
Usuário/Senha



Diálogo de confirmação
Aplicativo Android



Senhas descartáveis (OTP)
com Google Authenticator

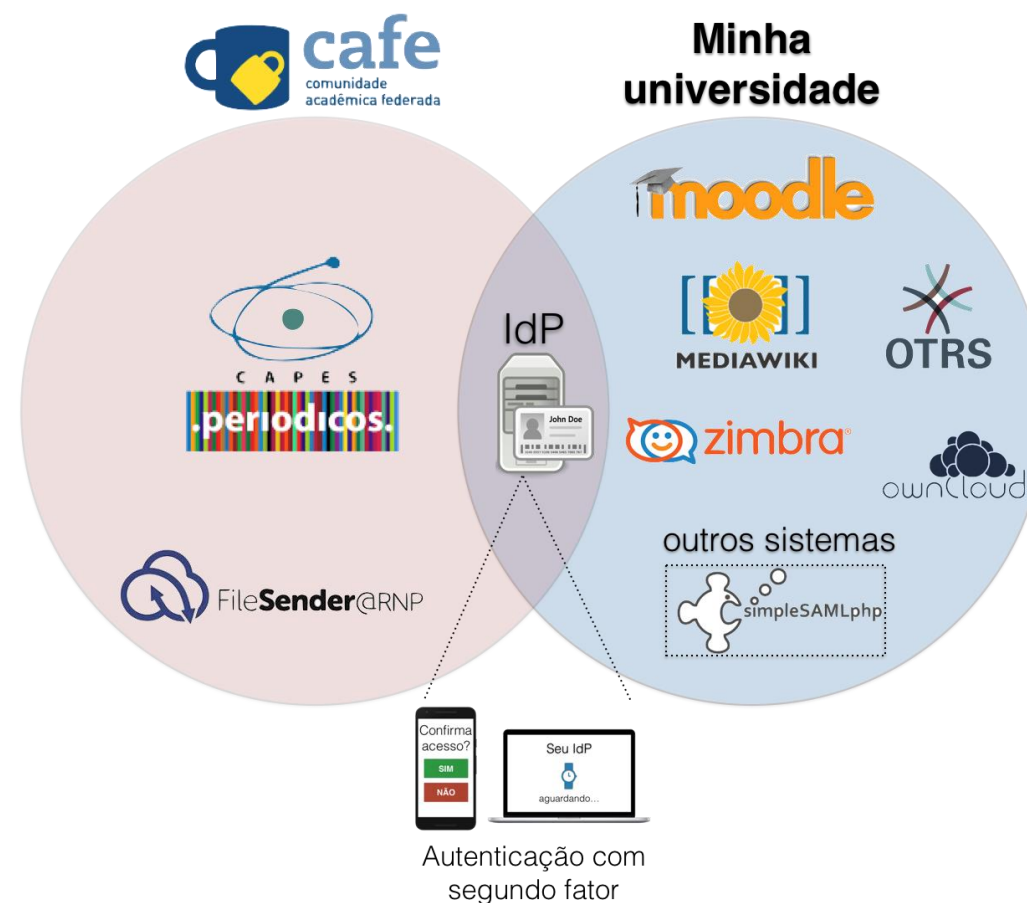


WebAuthN (FIDO2) – novo
padrão para autenticação
robusta na web

- **Como estes resultados podem contribuir para a construção de patrimônio digital, de dados abertos ou de tecnologias de memória?**
 - Provedores de Serviços da CAFe
 - Primeira solução completa e **de código aberto** que permite a autenticação multi-fator em provedores de identidade Shibboleth
 - **Construção modular** permite que instituições estendam a solução para inserir fatores de autenticação que sejam mais adequados ou já usados por estas

■ Qual a visão de futuro para aplicação dos resultados no curto prazo (6 a 18 meses)?

- Autenticação mais **robusta** na CAFe
- Instituições podem usar a solução para aumentar a robustez da autenticação em seus sistemas internos



Obrigada

Michelle Wingham

wingham@univali.br



1. RDP - Rede de Dados de Pesquisa Brasileira

Rafael da Rocha (UFRGS)

2. EUBra-BIGSEA Lemonade - Europe-Brazil Collaboration of Big Data Scientific Research through Cloud-Centric Applications

Walter dos Santos Filho (UFMG)

3. Athmosphere - Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing

Francisco Brasileiro (UFCG)

4. GT-IPÊ Analytics - Transformando dados brutos de monitoramento em informações valiosas ao gerenciamento da rede

Marinho Barcellos (UFRGS)

5. GT-AMPTo – Autenticação Multi-fator para Todos

Michelle Silva Wangham (UNIVALI)

Visitem os posters ao lado do estande da RNP!

Gratidão!

André Marins (RNP)

andre.marins@rnp.br