



The Tahopen Project – Universal Business Intelligence Platform

Marcio Junior Vieira
Team Leader - Tahopen
CEO & Data Scientist, Ambiente Livre
Pesquisador da UFG.

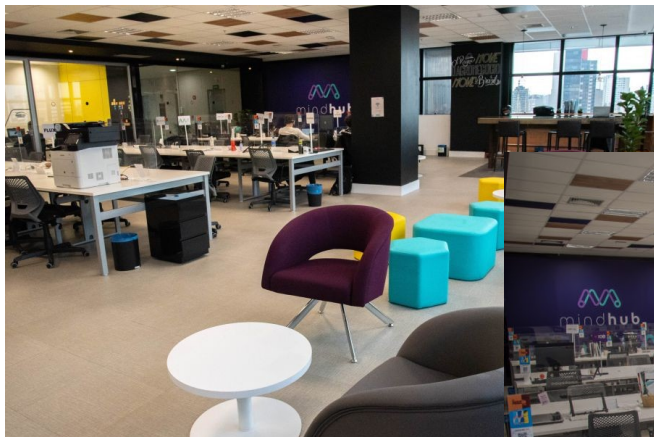
Mini-CV

- 25 anos de experiência em TI, vivência em desenvolvimento e análise de sistemas de gestão empresarial e ciência de dados.
- CEO da Ambiente Livre atuando como Cientista de Dados, Engenheiro de Dados e Arquiteto de Software.
- Professor dos MBAs em Big Data & Data Science, Inteligência Artificial e Business Intelligence e Analytics da Universidade Positivo.
- Professor do MBA Artificial Intelligence e Machine Learning da FIAP.
- Pesquisador do Laboratório de Tecnologias para Tomada de Decisão da Universidade de Brasília (Unb/Latitude).
- Trabalhando com Free Software e Open Source desde 2000 com serviços de consultoria e treinamento.
- Graduado em Tecnologia em Informática(2004) e pós-graduado em Software Livre(2005) ambos pela UFPR.
- Palestrante FLOSS em: FISL, TDC, Latinoware, Campus Party, Pentaho Day, Ticonova, PgDay e FTSL.
- Organizador Geral: Pentaho Day 2017, 2015, 2019 e apoio nas ed. 2013 e 2014.
- Data Scientist, instrutor e consultor de Big Data e Data Science com tecnologias abertas.
- Ajudou a capacitar equipes de Big Data na IBM, Accenture, Tivit, Serpro, Natura, MP, Netshoes, Embraer entre outras.
- Especialista em implantação e customização de Big Data com Hadoop, Spark, Pentaho, Cassandra e MongoDB.
- Contribuidor de projetos internacionais, tais como Pentaho, LimeSurvey, SuiteCRM e Camunda.
- Especialista em implantação e customização de ECM com Alfresco e BPM com Activiti, Flowable e Camunda.
- Certificado (Certified Pentaho Solutions) pela Hitachi Vantara (Pentaho).
- Membro da The Order Of de Bee (comunidade Alfresco para desenvolver o ecossistema Alfresco independente)
- Trabalha profissionalmente com Apache Zeppelin desde 2019.



Open Software for Business

- Fundada em 2004 com foco em consultoria com FLOSS.
- Experts em 34 soluções para geração de negócios com Software Livre/Código Aberto.
- Atualmente estamos sediados no Hub de Inovação Mindhub em Curitiba (FAE).



Nosso Ecossistema de Serviços

Big Data e Data Science	CRM e CMS	ECM e BPM	Business Intelligence
Análise de Dados da IoT Análise Preditiva Processamento Distribuído Banco de Dados Colunares Big Data & Data Lake Big Data Analytics Machine Learning Consultoria Treinamento Projeto	Marketing e Vendas Fidelização SAC e Pós-vendas Portais de Conteúdo Customer Relationship Management Content Management System Pesquisa de Mercado & SLA Consultoria Treinamento Projeto	Gestão de Documentos Gerenciamento de Mídias Processo de Negócio BPMN e BPMS Enterprise Content Management Records Management Business Process Management Consultoria Treinamento Projeto	Painéis de Indicadores Cubos de Análise Relatórios Gerenciais Tomada de Decisão Business Intelligence & Analytics Dashboards e OLAP Data Integration & Data Mining Consultoria Treinamento Projetos



Tahopen

- Plataforma de análise de negócios de código aberto, fork do Pentaho.
- Ele fornece uma solução completa para coleta, integração, análise e visualização de dados. Com uma interface amigável e ferramentas robustas,
- Permite que as organizações transformem grandes volumes de dados em insights acionáveis, melhorando os processos de tomada de decisão.
- Oferece suporte a processos ETL, relatórios dinâmicos, painéis interativos e análises preditivas, integrando-se perfeitamente a vários sistemas e fontes de dados.
- 100% Open Source.

About Tahopen

🌟 Announcing the Tahopen Project: An Open Source Alternative for the Pentaho Community! 🌟

Dear Pentaho Community,

We are excited to introduce the **Tahopen Project**, a fork of Pentaho version 10.1! As Pentaho transitions to a Business Source License with version 10.3, we recognize the need for an open and accessible alternative. Tahopen aims to keep the spirit of open-source alive, providing a platform for collaboration and innovation.

We invite all members of the Pentaho community to join us in this exciting endeavor. Your contributions, whether through code, ideas, or feedback, will be invaluable as we build a project that serves everyone.

If you're interested in collaborating, please visit our GitHub repository: [GitHub - Tahopen Project](#). Together, we can ensure that the benefits of Pentaho's capabilities remain accessible to all.

e.com.br



A Empresa

- A Hitachi Vantana é a atual desenvolvedora do Pentaho
- Resultado da Fusão da Pentaho Inc, HDS (Hitachi Data System) e Lumada (IOT)
- A Hitachi Vantana é uma empresa do Grupo Hitachi
-

HITACHI
Inspire the Next



Linha do Tempo

- A Pentaho Inc foi fundada em 2005
- A mesma integrou diversos produtos open source e contratou diversos lideres dos projetos e em seguida lançou a Plataforma Pentaho BI
- Criou uma versão CE (Community Edition) e uma EE (Enterprise Edition)
- Vem aprimorando a plataforma (Tanto a versão CE como EE)
- Em 2013 a Pentaho adquiriu a empresa WebDetails (Desenvolvedora das Ctools)
- Em 2015 a Pentaho foi adquirida pela Hitachi e incorporada a Hitachi Data System mas se manteve totalmente independente
- 09/2017 anunciada a criação da Hitachi Vantara fusão de 3 empresa do grupo
- 2017 Lançada a Plataforma Pentaho 8.0
- Atualmente a sede esta na Florida – EUA

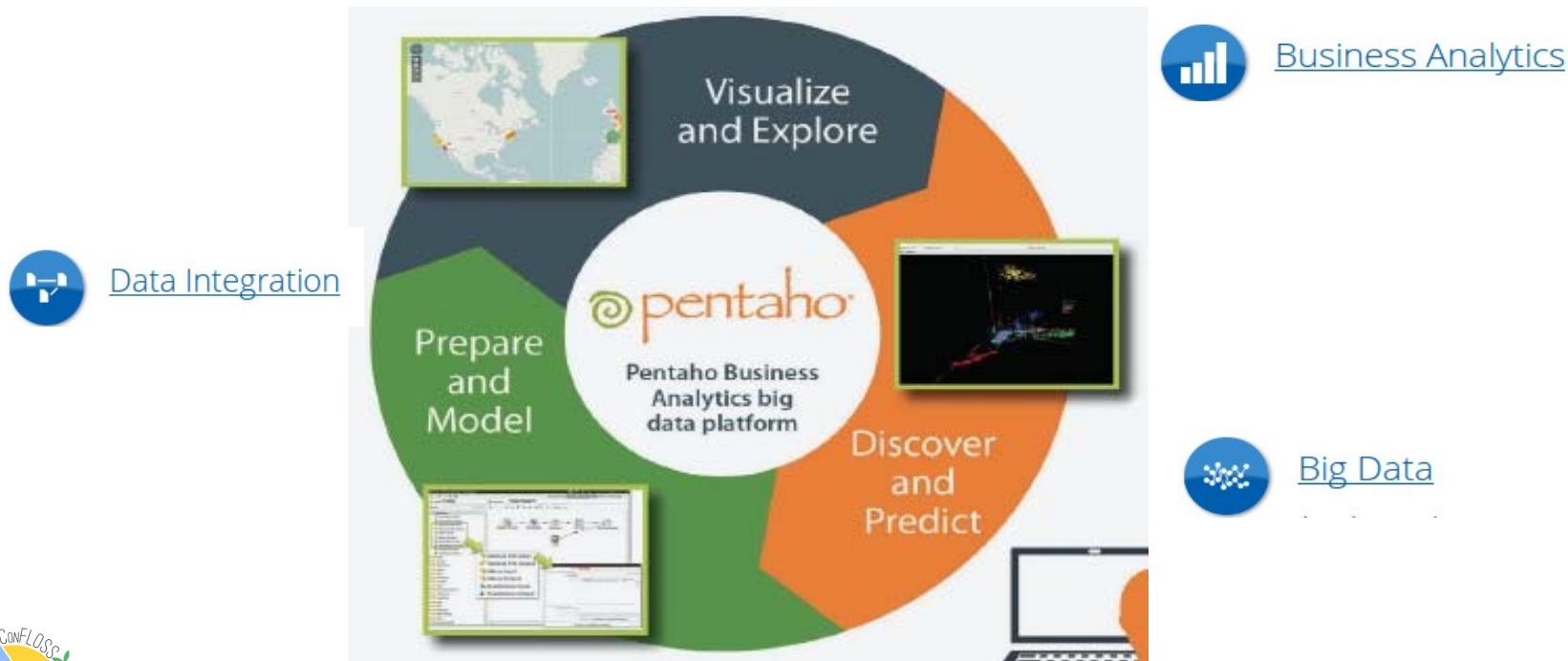


HITACHI
Inspire the Next



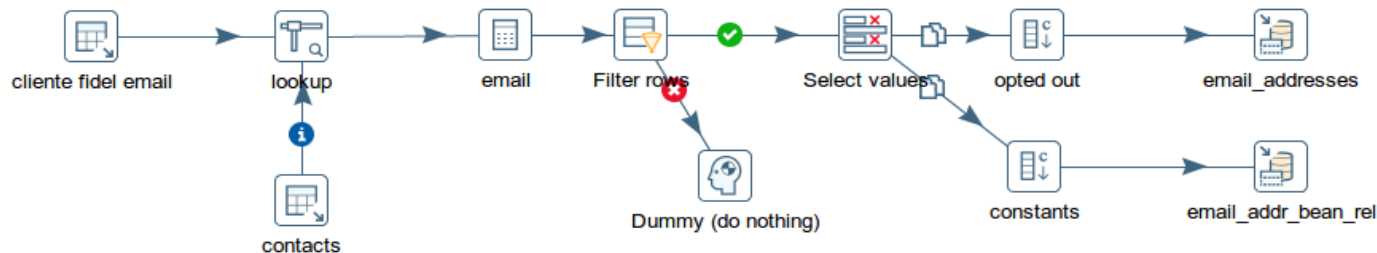
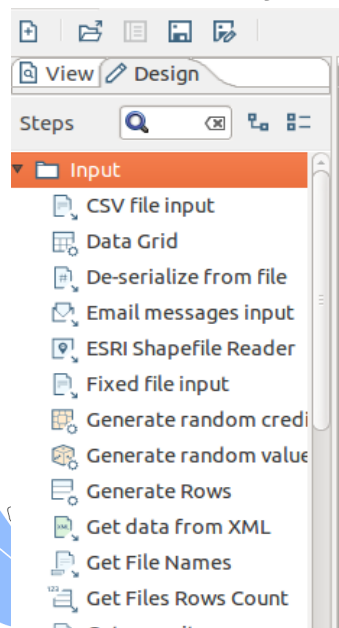
3 Pilares do Pentaho

- Plataforma abrangente para integração de dados e Business Analytics.

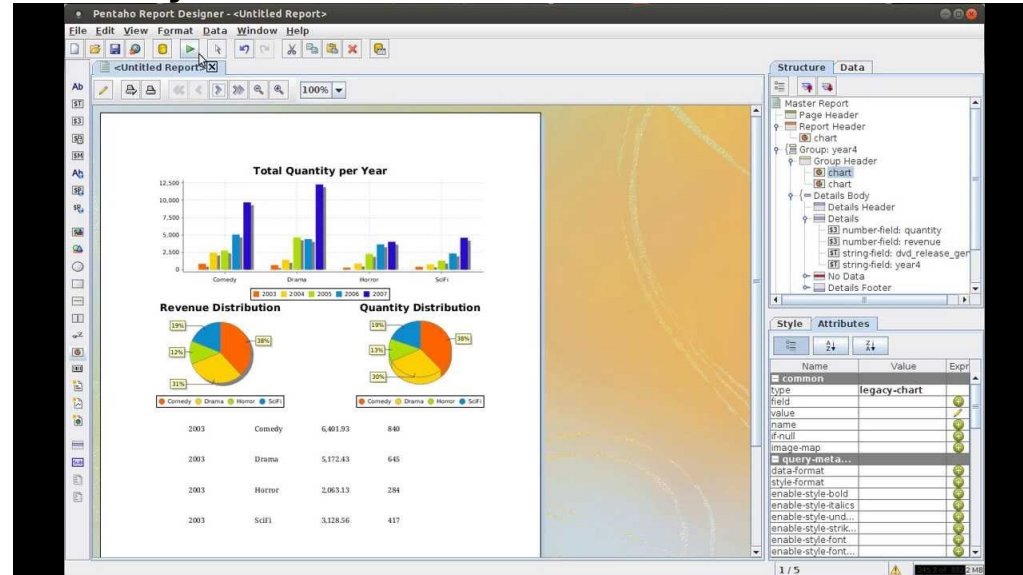


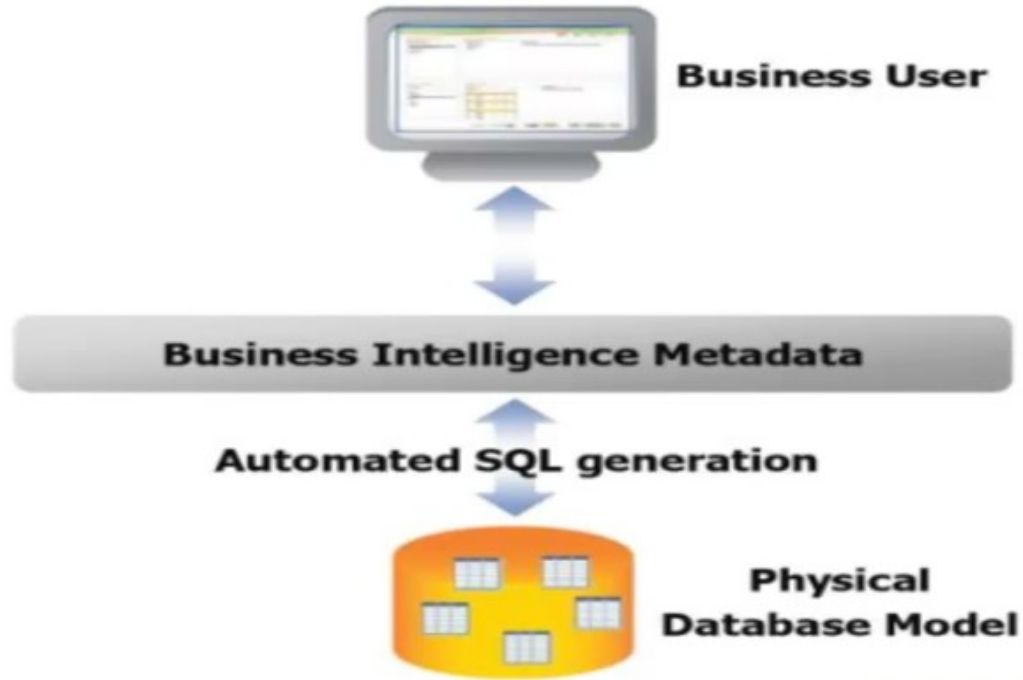
Pentaho Data Integration (PDI)

- Processa em Paralelo (Também roda em Cluster Spark)
- Acessar dados diretamente (se necessário sem DW)
- Permite publicar dados diretamente em Reports, Ad-Hoc Reports e Dashboards.
- “Programação e Fluxo Visual” com aproximadamente 350 steps diferentes



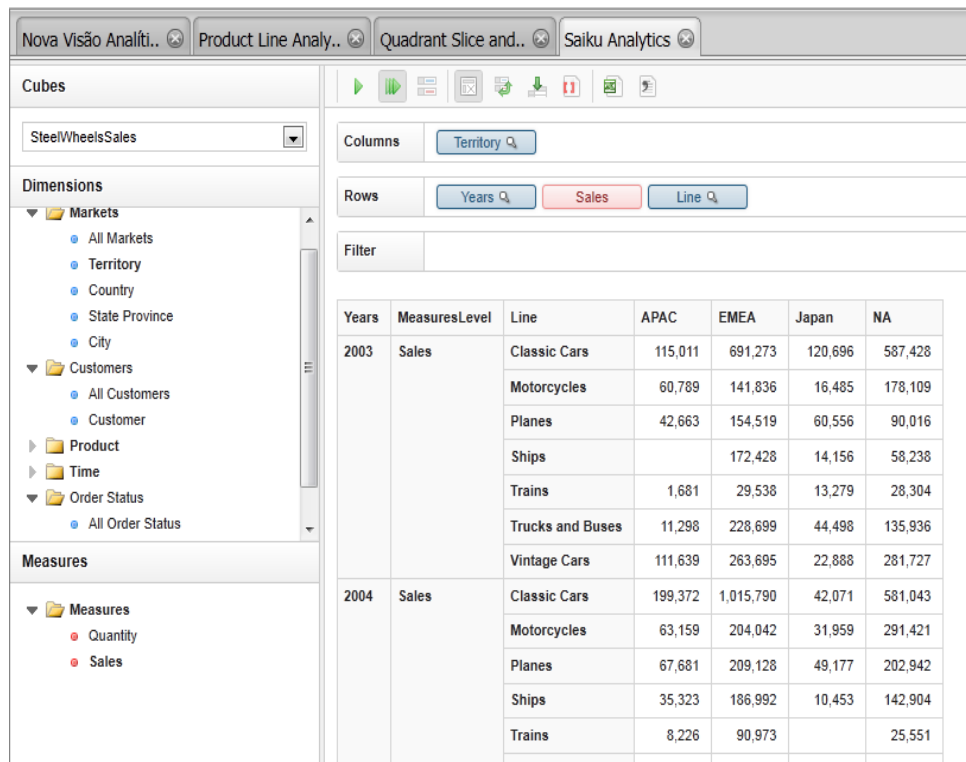
- Visualização Web ou Embed.
- Assistente de geração de relatórios
- Amplo suporte de fonte de dados, incluindo relacionais, OLAP, XML e Pentaho Analysis, arquivos flat, objetos Java e ...
- **Big Data Reports (integra-se com PDI)**





Pentaho Analysis (Mondrian)

- Explorar livremente através de perfuração e tabulação de dados.
- Velocidade de resposta as consultas analíticas complexas.
- Visualiza a informação multi-dimensional, escolhendo métricas e atributos específicos para analisar.
- Totalmente Web, drag-and-drop.
- Classificação e filtragem avançadas
- Totais personalizados e definidos pelo usuário cálculos
- Gráfico para visualizações



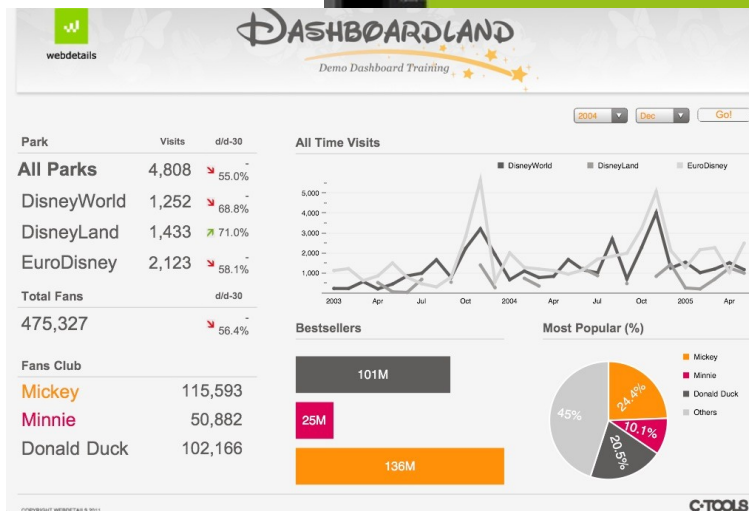
The screenshot shows the Pentaho Analysis tool interface. The left sidebar contains a tree view of dimensions and measures. The main area displays a table of sales data for SteelWheelsSales, filtered by Territory, Years, Sales, and Line. The table shows data for 2003 and 2004, categorized by Line (Classic Cars, Motorcycles, Planes, Ships, Trains, Trucks and Buses, Vintage Cars) and further broken down by region (APAC, EMEA, Japan, NA).

Years	MeasuresLevel	Line	APAC	EMEA	Japan	NA
2003	Sales	Classic Cars	115,011	691,273	120,696	587,428
		Motorcycles	60,789	141,836	16,485	178,109
		Planes	42,663	154,519	60,556	90,016
		Ships		172,428	14,156	58,238
		Trains	1,681	29,538	13,279	28,304
		Trucks and Buses	11,298	228,699	44,498	135,936
		Vintage Cars	111,639	263,695	22,888	281,727
2004	Sales	Classic Cars	199,372	1,015,790	42,071	581,043
		Motorcycles	63,159	204,042	31,959	291,421
		Planes	67,681	209,128	49,177	202,942
		Ships	35,323	186,992	10,453	142,904
		Trains	8,226	90,973		25,551



Dashboards

- Oferecer métricas em uma atrativa interface intuitiva e visual, painéis proporcionam aos gestores empresariais a informação crítica de que necessitam para entender e melhorar o desempenho organizacional.



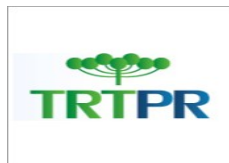
Pentaho Day



2019
PentahoDay
Curitiba - Brasil
10 e 11 de Maio



Empresas Brasileiras Usando Pentaho



Empresas Usando Pentaho



BSL

- A partir da versão 10.3, a Pentaho Community Edition, que era a versão de código aberto, foi renomeada para Developer Edition e seu modelo de licença foi alterado para a Business Source License (BSL).

- **BSL Não é Open Source!**

- A BSL da Hitachi Vantara para o Pentaho 10.x estipula que a licença mudará para a Apache License 2.0 quatro anos após a data de lançamento da versão específica.
- O Pentaho 10.3, lançado em 2024, se tornará de código aberto em 2028

Será ?



Sourceforge.net

- **Sempre tenha seus fontes!**
- **Hitachi repo mavens descontinuados...**

legacy-pentaho-ce

Legacy versions - Pentaho CE

Sourceforge.net

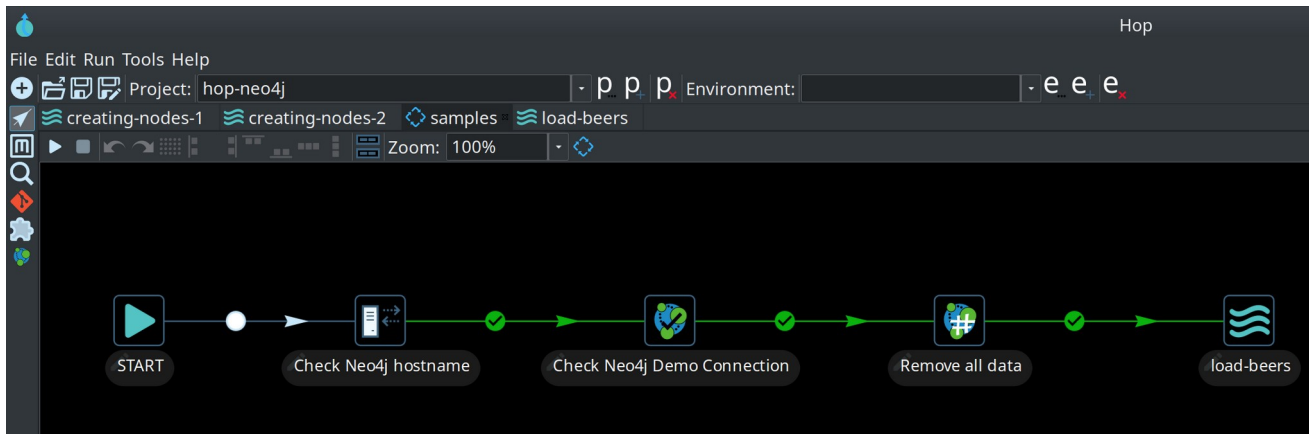
Older versions of pentaho have been removed from sourceforge.net/projects/pentaho , let's group some versions here for the use case in legacy projects.

For official Hitachi downloads go to: <https://www.hitachivantara.com/en-us/products/pentaho-plus-platform/data-integration-analytics/pentaho-community-edition.html> Currently only versions 9.4, 9.3, 9.2 and 8.3.



Acrônimo: Hop Orquestration Plataform

- Orquestração
 - **Dados** – Pipelines e Workflows.
 - **Metadata** – Edição, Manuseio e gerenciamento.
 - **Insights**: Execução e tratamento de dados, log do processo.
- - **Configurações**: Manuseio de ecossistemas complexos.



Background

- Iniciativa de uma Comunidade (junto com Matt Casters).
- Fork do Kettle/Pentaho/PDI 8.2 + WebSpoon + Patches + plugins...
- (Somados são mais de 20 anos de desenvolvimento de software)
- Interface gráfica renovada.
- Back-end de metadados novo.
- Maior simplicidade.
- Muitos códigos refatorados.
- Licença Apache v2.0



Contributors 28



+ 17 contributors



Tahopen

- Pentaho - Tahopen
- PDI → TDI
- Mondrian → Driamon
- Pentaho Dashboard – Tanhopen Dashboard
- PSW → TSW
- Saiku – Kusai
- PME - TME



Tahopen

- Pentaho 10.1 – Tahopen 0.1
- Atualmente temos já a Tahopen 0.2 em build

Downloads

Select the version of Tahopen to download:

Beta

- [Tahopen Schema Workbench 0.1.0.0](#)

Snapshot

- [Tahopen Reporting Designer 0.1.0.0 \(SNAPSHOT\)](#)
- [Tahopen Metadata Editor 0.1.0.0 \(SNAPSHOT\)](#)

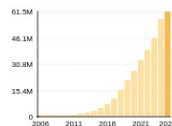


Tahopen Fork Status

Tahopen

- 101 repositórios clonados da Pentaho
- 59 Convertidos e Compilados na 0.1
- 16 packages publicados já em Maven (Maven central)

Indexed Artifacts (61.5M)



Popular Categories

Testing Frameworks & Tools

Android Packages

JVM Languages

Logging Frameworks

Java Specifications

JSON Libraries

Core Utilities

Mocking

Web Assets

Annotation Libraries

Language Runtime

HTTP Clients

Logging Bridges

Dependency Injection

XML Processing

Web Frameworks

Home » org.tahopen

Group: Tahopen

Sort by: Popular ▼



1. **Tahopen Commons Connections**

org.tahopen » tahopen-connections

1 usages

Apache

The Tahopen Connections API defines a common set of interfaces for dealing with connections and result sets.

Last Release on Jan 3, 2025



2. **Tahopen CWM**

org.tahopen » tahopen-cwm

LGPL

Common Warehouse Metamodel.

Last Release on Jan 11, 2025



3. **Commons Database Model**

org.tahopen » commons-database-model

LGPL

Commons Database Model.

Last Release on Jan 4, 2025



4. **Tahopen Community Edition Project Parent Pom**

org.tahopen » tahopen-ce-parent-pom

Apache

a parent pom for general configuration of Tahopen sponsored open source projects

Last Release on Dec 22, 2024



Tahopen

- Apache 2 na maioria dos componentes
- GPL em algumas Libs.



Tahopen

- Dez/2026 – Lançar a versão 0.2
- Submeter Projeto para alguma fundação (Ex. Apache, CNCF, etc)
- Empacotar em .deb (Debian packages)
- Remover Data Integration e Incorporar Apache Hop;
- Distribuir IAC (Docker , Docker-compose, K8s, Helm, etc)
- Desenvolver Kubernetes Operator



Tahopen

- Converter repos, remoção de bibliotecas do Repo Hitachi
- Atualização de Libs
- Temos 100 repos escolha 1 e ajude!
- Documentação Para novos devs
- Site

Requisitos Básicos

- Java
- Apache Maven
- Para Documentação: UX e Markdown
- Para IAC : DevOps



Vamos Começar Já! DEMO! Fork!

Tahopen

- `git clone https://github.com/tahopen/tahopen-aggdesigner`

CODE... CODE... CODE... CODE...

IA CODE...

CODE... CODE... CODE... CODE...

- Build
`mvn clean install`
- Push
`git push`



- **Documentação Tahopen**
<https://www.tahopen.org>
- **Fontes GitHub**
<https://github.com/orgs/tahopen/>



Obrigado

Marcio Junior Vieira

marcio@ambientelivre.com.br

@marviojvieira @ambientelivre

<https://www.linkedin.com/in/mvieira1/>

Blog: <http://blogs.ambientelivre.com.br/marcio/>

Slides:

<https://github.com/ambientelivre/labs>