

Wendelin Big Data Analytics

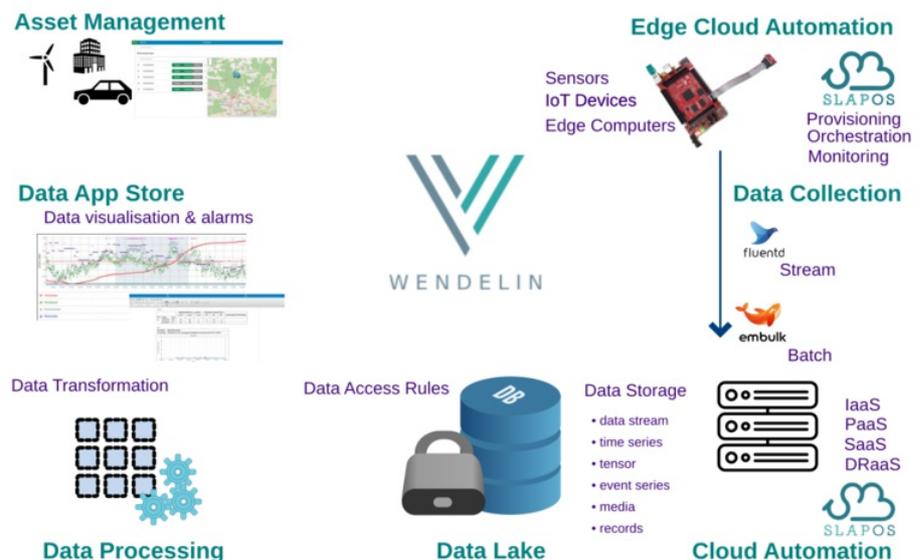


Big Data Analytics, Monitoring and Machine Learning Using Wendelin

Wendelin is an open-source, 100% Python-based platform for data ingestion, storage, analysis, and visualisation. Together with Nexedi's edge and cloud computing system SlapOS, Wendelin industrialises the complete chain of data-based applications from the collection of data at the edge over local IoT bridges to the central data analysis and visualisation on the cloud. Thanks to the integration of popular data science technologies such as Jupyter, SciPy, Pandas, Plotly and scikit-learn, the Wendelin platform provides an extensive collection of data analysis tools.

Key Benefits

- Ingestion, Storage, Analytics, Visualisation
- Cost-Reduction through Open-Source
- Independence: Public/ Private Cloud or on-Premise
- Unification: Python for Analytics and Cluster Environment



Scenario:

Data Lake

- Data Lake for Batch and Streaming Data
- Distributed Transactional Object Storage
- Plugins for >100 types of Data Sources
- Process-orientated Metadata Management
- Dynamic Data Access Rules

Wendelin as an Alternative to AWS and Palantir

Wendelin provides a cost-efficient alternative to conventional public clouds (e.g. AWS, Azure) and proprietary AI platforms (e.g. Palantir). Wendelin clusters can be provisioned on public as well as on private clouds or on-premise avoiding lock-in to a single vendor.

Wendelin is 100% open-source, patent-free and developed in Europe. Thus it has an advantage over US-American clouds and AI platforms when it comes to convincing clients or suppliers to ingest their sensitive data into a Wendelin data lake. Furthermore, Wendelin does not have any restrictions to be used in countries such as China (unlike other solutions due to US export regulations).

Scenario:

Data Science Industrialisation

- Automation of Edge and Cloud Services
- Data Processing Workflow Automation
- Native Python out-of-core Processing
- Scalable Visualisation



©Nexedi 2020

Nexedi
147 rue du ballon
59110 La Madeleine
France
info@nexedi.com
+33(0)-3-59 05 15 26

Printed in France
2020-Jun
All rights reserved

All other company, product, or service names may be trademarks or service marks of others and are the property of their respective owners. References in this publication to the companies products or services do not imply that the company intends to make these available in all countries in which it operates.

The customer is responsible for ensuring compliance with legal requirements. It is the responsibility of the customer to seek the advice of competent legal counsel as to the identification and interpretation of relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may have to take to comply with these laws.



Application Scenarios

Wendelin is suited for industrialisation and unification of data-based services from data collection to visualisation. The ingestion and transformation of data is modelled by a business process-oriented approach, enabling traceability of data supply chains and pricing of data transformations.

Data Lake

Wendelin supports the real-time collection of data from multiple sources (machines, websites, e-commerce, clients and suppliers). Data engineers can choose from more than 100 ready-made plugins for different web services and databases thanks to the integration of ARM's open-source data collection solutions Fluentd (for streaming) and Embulk (for batch data). Collected raw data can then be aggregated and structured with PyData libraries such as Pandas and SciPy and finally be analysed automatically with machine learning tools such as scikit-learn or TensorFlow. Examples are intelligent searches and finding correlations.

Data Science Industrialisation

The business process-oriented approach of managing data analysis operations makes Wendelin a perfect fit for unification and automation of recurring data science tasks on a production system. Wendelin unifies data engineering through the usage of Python on both the analysis environment and the production cluster. All data is stored in Wendelin's distributed transactional object database NEO. Native out-of-core access to persistent NumPy ndarrays provided by the `wendelin.core` library allows for scalable analytics. Analysis operations can be implemented without restrictions on the available memory, and they do not need to be recompiled for running on the production cluster. With `wendelin.core` the complete NumPy API is available when accessing out-of-core ndarrays (unlike other solutions such as Apache Spark or Dask, which depend on a compatibility layer).

Success Cases & Services

The components of the Wendelin platform, ERP5, NEO and SlapOS are successfully used in automotive (PSA, Toyota) and aerospace industries (Airbus Defence and Space). Wendelin itself is used in Germany for ice detection on wind turbines and noise and vibration monitoring of construction sites. Services provided by Nexedi cover automation of data collection at the edge, deployment and management of a Wendelin cloud, analysis and visualisation as well as consulting on single topics or implementation of full-fledged big data applications on the Wendelin stack.