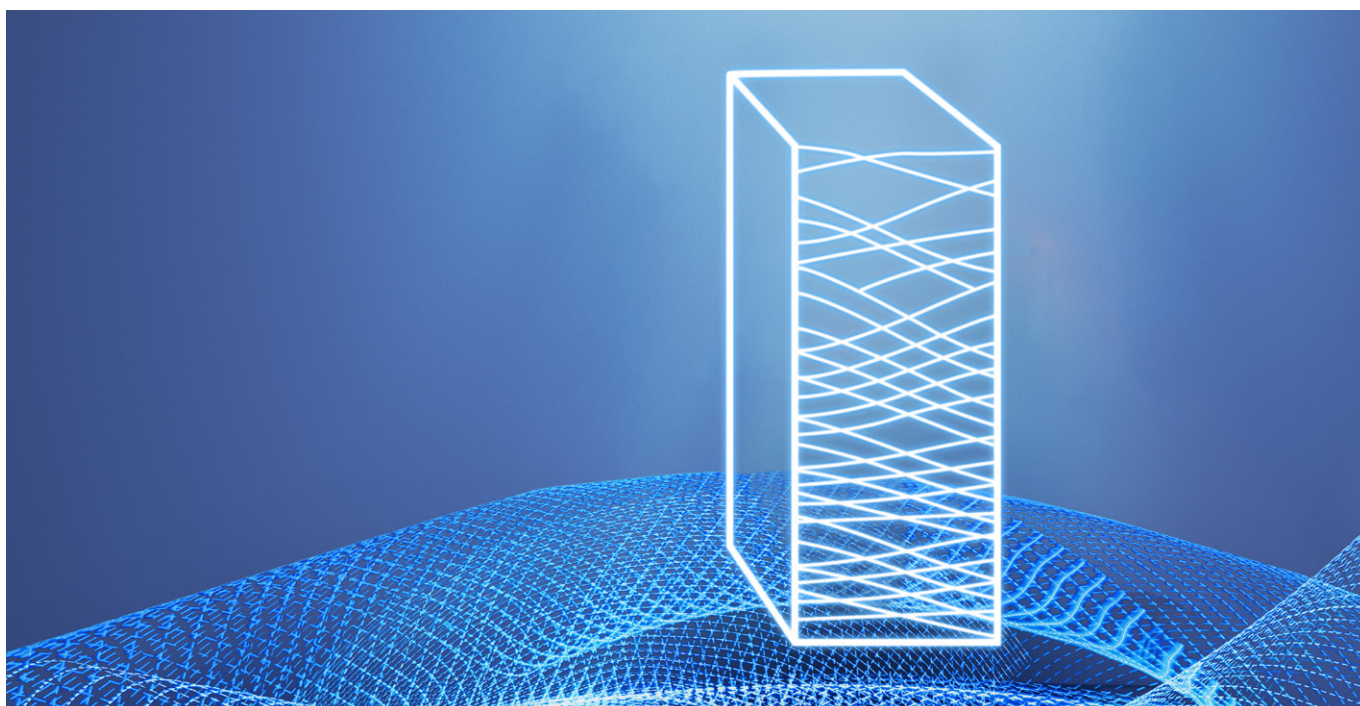


Trust your data



“Though 85% of companies are trying to be data-driven, only 37% of that number say they’ve been successful”

“About three out of five people in leadership roles say a failure to get on board with big data could lead to obsolescence”

As of this moment, only 0.5% of all accessible data is analyzed and used. Imagine the potential here.

Building trust in your data

Digital transformation is a business requirement for organizations. Data is critical in executing this transformation, enabling innovation and competitive differentiation. The results for data-driven organizations are better decision-making abilities with an immediate positive impact.

While data lakes are being used by enterprises, questions and challenges are arising:

- Who in my organization is using data and how?
- Is my data secure and can I trust it?
- How do we handle large quantities of structured and unstructured data?

Data-driven digital transformation requires robust and suitable data architecture with strong governance not only to maintain security but also to enable activities such as analytics, enterprise artificial intelligence and interconnected objects.

Here is where Codex Datalake Engine offers the solution. Meeting your trust and compliance needs and leaving you free to get maximum value from your data.

Codex Datalake Engine is a well-defined, new generation of data lake solution. It is an end-to-end data management and security platform, which enables organizations to build a highly scalable, easy-to-use and cost-effective private cloud solution. It is certified by Cloudera.

Enabling visibility and control of data

Codex Datalake Engine addresses the needs of all stakeholders interacting with data: Chief data officers, data stewards, compliance groups and data scientists, AI developers and BI users.

Chief data officers, data stewards have the following concerns:

- How is data being used?
- How can data be optimized?
- How do you efficiently manage the lifecycle of your data?
- How do you overcome data silos?
- How do you ensure data regulations are met?

Compliance groups track and protect access to sensitive data. Their primary task is to always be prepared for an audit: tracking who is accessing data, what data they're accessing and what it is being used for. Their job is to ensure that sensitive data is well governed and protected and in line with General Data Protection Regulation (GDPR) within the European Union and the European Economic Area.

Data scientists, AI developers, BI users need to find the data that matters the most for their business. They want to be able to explore data, trust what they find, and be able to visualize relationships between data sets and make the most from their data.



Ensuring data quality

Codex Datalake Engine is a consistent data management and security appliance, providing a data governance solution and the management of the complete data life-cycle, from data ingestion, data cleansing, data blending, data discovery, audit, data lineage and policy enforcement. Codex Datalake Engine improves and ensures data quality and makes data trustworthy.

Bringing data security and protection on-premises

According to the Identity Theft Resource Center, 2016 saw 1,093 data breaches, a 40% increase from 2015. 2017 almost hit that mark by July.

Organizations are very concerned with data security and protection. With Codex Datalake Engine, all data is stored on site and organizations keep full control over their data and its lifecycle as well as full control over the infrastructure, the applications and the operations, enabling more compatibility and thus minimizing the risk of failure. The complete virtualization minimizes the configuration costs and avoids extra cost when introducing new applications.

Running a cost-effective, scalable and easy-to-use datalake on site

Codex Datalake Engine is a preconfigured, scalable, easy-to-use and fully virtualized appliance. It is cost-effective with minimal administration needs or energy usage. As a result, organizations spend less time installing, tuning, operating, troubleshooting, patching, upgrading, and dealing with integration, adoption of technologies and scale-related issues.

Codex Datalake Engine can also be the foundation for data in a private cloud.

An appliance certified by Cloudera

Cloudera provides a scalable, flexible, integrated platform that makes it easy to manage rapidly increasing volumes and varieties of data. Cloudera products and solutions enable to deploy and manage Apache Hadoop and related projects, manipulate and analyze data, and keep that data secure and protected.

Codex Datalake Engine is Cloudera certified. It is the result of a joint effort between Atos and Cloudera to achieve a solid and trustworthy data architecture and deliver the most complete, secure, industrial and qualified datalake solution on the market.

Leveraging the power of the ultimate BullSequana S server

Codex Datalake Engine is delivered as an appliance that features BullSequana S, one of the most agile, scalable and open servers. With its dynamic reconfiguration capabilities, BullSequana S server combines exceptional performance with unprecedented levels of agility and efficiency. The BullSequana S server range features from 2 to 32 Intel® Xeon® Scalable processors in a single server.

Full power in a virtualized and cost-effective design

All components of Codex Datalake Engine are integrated into a compact, cost optimized platform. For this, Codex Datalake Engine takes advantage of virtualization, which allows it to optimize the hardware sizing needed for each software component, instead of using oversized separate machines.

Codex Datalake Engine for development

For development and test purposes Codex Datalake Engine is bundled into a single compact server, scaling from BullSequana S200 (2 CPUs) hosting 5 TB of user data, up to BullSequana S800 (8 CPUs) hosting 20 TB to 50 TB of user data.

Codex Datalake Engine for production

For production configuration starts with 3 BullSequana S200 and can be extended up to 12 BullSequana S200, with 50TB to 350 TB of user data. For hefty workloads, Codex Datalake Engine leverages the BullSequana S400, starts with 6 BullSequana S400 and can support several PB of user data with 60 TB upgrade increments

Codex Datalake Engine key functionalities

Searching and visualizing data lineage:	It enables chief data officers, business analysts and data scientists to define, search, update the properties, and tag data entities as well as to view relationships between datasets.
Metadata:	It makes data more accessible and makes it easy to extract value from the data lake. Metadata provides significant information about the background and significance of the data stored in the data lake to its users. This is accomplished by intelligently tagging every bit of data as it is ingested.
Auditing data access and controlling access privileges:	Auditing features add secured, real-time audit components to key data and access frameworks. This allows compliance groups to configure, collect and view audit events and understand who accessed what data and how.
Policies:	Policy features enable users to specify automated actions based on data access or on a schedule to create alerts or remove data for example.
Analytics:	Its analytics features enable Hadoop administrators and business analysts to examine data usage patterns and create policies based on those patterns.
Data encryption and key management:	It provides a critical layer of protection against potential threats. Encryption and key management are also required for meeting key compliance initiatives and ensuring the integrity of enterprise data. It transparently encrypts and secures data without requiring changes to applications and ensures there is minimal performance lag in the encryption or decryption process. It uses a trustee server as enterprise grade virtual safe-deposit box to store and manage cryptographic keys.

Empowering the deployment of Big Data and AI applications

Codex Datalake Engine is the foundation to develop Big Data and AI applications. By leveraging all available data across organizations, it gives a reliable and comprehensive view of data. Analytics and AI developers have a full visibility and control of all data.

Codex AI Suite, the new AI software suite, designed to accelerate the development of AI applications, leverages Codex Datalake Engine to benefit from a complete data aggregation and achieve the highest degree of accuracy. Codex AI Suite can handle data sets at scale, store results and secure data on-premises.

Codex Datalake Engine is a key component of Codex

Atos Codex is a portfolio of services and products designed to accelerate and industrialize the enterprise adoption of the data analytics and related activities central to all serious digital transformation strategies.

Atos Codex spans data analytics, AI, deep learning and machine learning, cognitive computing and the Internet-of-Things.

Atos Datalake Engine is a key component of Codex when it is about enabling organizations to become data-driven. Codex is providing a complete set of services to support organizations to become data-driven. It includes several strategic steps:

- Data governance to define the strategy, the roles, responsibilities and the operating model
- Data organization to define who are the decision makers, i.e. chief data officers, data stewards, data scientists
- Data Landscape to define the standards and where and how data is used
- Data process to properly manage and maintain the quality of the data
- Technology to design the datalake and the tools to support the data governance, data organization, data landscape and data processes.

About Atos

Atos is a global leader in digital transformation with approximately 100,000 employees in 73 countries and annual revenue of around € 12 billion. European number one in Big Data, Cybersecurity, High Performance Computing and Digital Workplace, the Group provides Cloud services, Infrastructure & Data Management, Business & Platform solutions, as well as transactional services through Worldline, the European leader in the payment industry. With its cutting-edge technologies, digital expertise and industry knowledge, Atos supports the digital transformation of its clients across various business sectors: Defense, Financial Services, Health, Manufacturing, Media, Energy & Utilities, Public sector, Retail, Telecommunications and Transportation. The Group is the Worldwide Information Technology Partner for the Olympic & Paralympic Games and operates under the brands Atos, Atos Consulting, Atos Worldgrid, Bull, Canopy, Unify and Worldline. Atos SE (Societas Europaea) is listed on the CAC40 Paris stock index.

Find out more about us

atos.net

Let's start a discussion together

