Functional description

Momenttum.ai

NTT DATA Business Solutions - Smart X

21 October 2024 | Document Version 1.3

Momenttum.ai

Contents

List of abbreviations

Abbreviation	Meaning
Momenttum.ai	The SaaS platform which this document is about.
SaaS	Software a as Service
AI & ML	Artificial Intelligence and Machine Learning
GenAI	Generative Artificial Intelligence
xxxOps	Ops refers to Operations

Referenced documents

Name	Content
Data Ethics	Document that describes our ethical codex when working with data
Data Integrity by Design	Document that describes how we managed, govern and work with data
Data Science Ethics	Document that describes our ethical codex when using data for machine learning
Model Complexity Classification	Document to help understand the impacts of complex machine learning and AI models

Document history

Date	Comments	Author
October 2024	Latest revision	Jonas Holck

1. Purpose of Momenttum

NTT DATA has developed momenttum.ai, a managed cloud computing service designed for advanced use cases. This platform aims to integrate digital and physical spaces, leveraging advanced technologies to meet human needs. Momenttum.ai is positioned to enhance quality of life and address various social challenges by utilizing data as a source of value and creativity through artificial intelligence, robotics, and biotechnology.

1.1. Capabilities

Receiving data

The infrastructure is based on Microsoft Azure components. It consists of 5 layers. The first layer is the landing zone in which all data sources can be handled either via an API-integration or via an IoT-hub. The next layer is the Data Warehouse in which the platform leverages an Azure-deployed Snowflake Database. From the Data Warehouse the data can be either used in the AI or BI platform for analysis and modelling to be presented lastly in the User Portal. The infrastructure leverages components such as API Management, Data Factory, Storage Accounts etc. for data handling and storage.

The platform can receive data from various sources, such as cloud services or databases, via APIs or Extraction methods. The data is then stored in the Landing Zone, which is a secure and scalable environment for data ingestion.

Data can be pushed to our API's, or it can be pulled from customer systems. Data can be extracted from, but not limited to:

Cloud sources

- Azure Blob Storage
- Azure Cosmos DB for NoSQL, PostgreSQL, and MongoDB
- Azure Data Explorer
- Azure Data Lake Storage Gen1 and Gen2
- Azure SQL Database and SQL Managed Instance
- Azure Synapse Analytics
- Amazon S3 and Redshift
- Snowflake
- Google BigQuery and Cloud Storage
- FTP, SFTP, and HTTP/S
- Numerous SaaS applications like Salesforce, Marketo, and ServiceNow

On premise using an integration runtime middleware

- Databases like SQL Server, Oracle, MySQL, PostgreSQL and DB2.
- NoSQL data stores such as MongoDB and Cassandra.
- Big data stores like HDFS, HBase, and Hive.

Extracting and cataloguing data

From the landing zone data is extracted and managed. The platform can extract metadata from the stored data and catalogue it in the Data Catalogue, which is a centralized repository for data discovery and governance.

Processing data

The platform can process the data using up 1024 CPUs this means that even very large datasets can be handled and processed quickly or in real time.

Managing and transformation of data

The platform can manage data using DataOps CI/CD, which is a framework for continuous integration and continuous delivery of data products. The platform can also convert the data into useful information by developing use cases for various applications, such as: Forecast, simulation, monitoring, dashboarding and digital assistants.

Search and Indexing

Search provides a powerful indexing capability that allows us to pull data from various data sources into a search index. This process is facilitated by an indexer, which is essentially a crawler that automates the data import and indexing process. In addition, we also deploy crawlers to fetch information from website to keep the index updated.

Forecasting and simulation

The platform can use the processed data to predict future outcomes and scenarios based on historical trends and patterns. The Data Science ML Ops includes language models and an analytics platform for natural language processing and data analysis.

Monitoring and dashboarding

The platform can use processed data to track and visualize key performance indicators and metrics in real-time.

Communication and Digital Assistant

The platform can use the processed data to send notifications and alerts to relevant stakeholders and users based on predefined rules and triggers. Data can be used be digital assistant to enhance and augment processes and systems.

1.2. Core Concepts

Energy saving Green Platform

NTT DATA has in collaboration with the Green Software foundation developed a few principals for green software. These principles implemented to Momenttum – which means the carbon footprint is actively reduced and processing only occurs when needed.

Read more here. https://learn.greensoftware.foundation/

Security by Design

We used Security by Design because it helps to ensure the safety and privacy of customer data and users in a world where cyber threats are constantly evolving and increasing. By building security into the design and development of Momenttum, Security by Design aims to prevent or mitigate potential vulnerabilities and risks, rather than relying on reactive measures after a breach or attack. Security by Design also helps to reduce the cost and complexity of maintaining and updating software systems, as well as enhancing user trust and satisfaction.

AI, ML, GenAI and Big Data capable

Momenttum includes capabilities to handle large amounts of data, and deploy, manage, and monitor ML Models and use AI's such as OpenAI's GPT and Meta's Llama models.

AI and ML models can be simple or complex, to understand what impacts this can create we have explained this in greater detail in the document "Model Complexity Classification".

Modular

Momenttum deploys are module-based design, this means that the capability of the platform can be extended and modified to accommodate for special customer needs.

IoT Integration

IoT smart devices and sensor are important sources of data, because of this Momenttum implements multiples mechanism to make best use of IoT sources. This includes connecting to sensors, receiving data and displaying and streaming data.

Tenants

Tenants are customers with a direct agreement with NTT DATA. Tenants are isolated and use their own resources for storing and managing data.

1.3. Data Lifecycle

Within the lifecycle of data, it is important to understand that different phases of the lifecycle serve different applications of that data. Both system users and machinery continuously generate data records, in order to run the business. These data records are then used in various follow-on processes in the form of actions, decisions, analytics, or other activities. For all of the follow-on processes it is really important that metadata that is created at the process level also is transferred to these follow-on processes.

Our default data life cycle is the following:

The details about how we process, manage, govern, and use data is described in the Data Integrity by Design document. This document describes the methods we use to extract value from data. The method is solid enough to be appliable in the rigid validation processes of the pharmaceutical industry, but agile enough when deploy with DataOps for less regulated industries.

1.4. Modules

The momenttum platform offers are number of core modules. Some of the are fully developed, and a few are still in an early stage and not fully integrated. Modules are shared across customer and updated as part of the platform release cycle.

Launchpad

This module handles access to resources, hierarchy of organisations and assignment of resources.

Reports

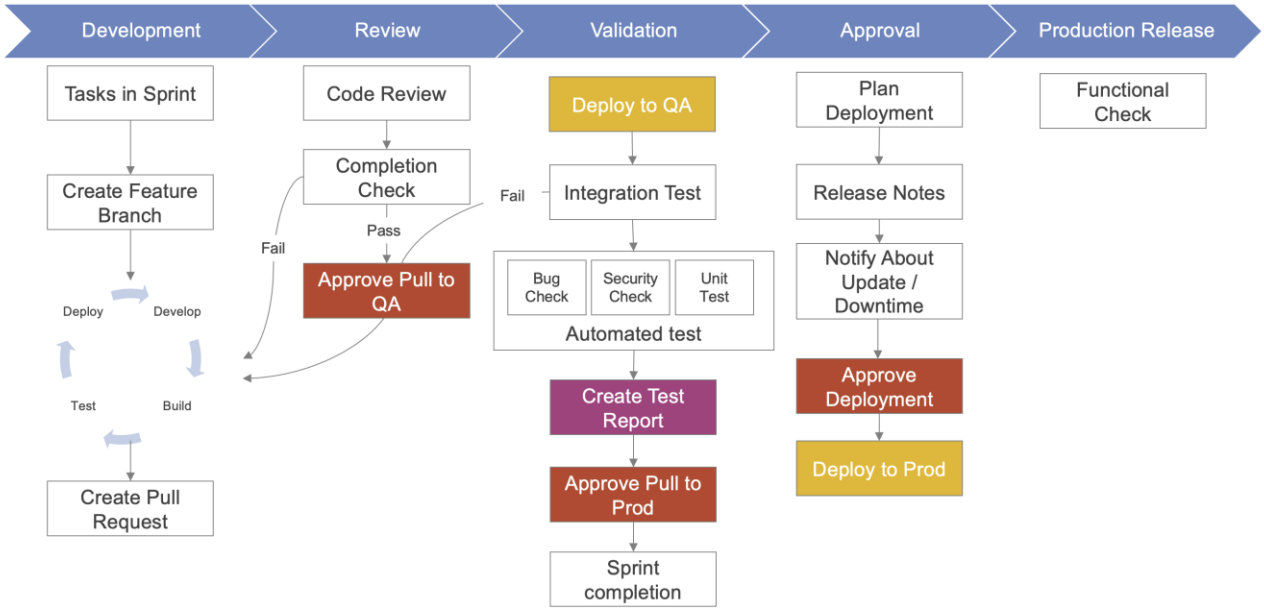
This module provides access to report and dashboard published in Momenttum.

1.5. Development Process

The development of new functionality is performed using DevOps methods. It's goes through the phases described below and is performed in 2-week sprint. The principle for four eyes applies to phases prior to review, validation, and production release.

Release is grouped into 2 categories: Customer and Platform. Customer releases affect only the customer and can be deployed at customer request. Platform releases are deployed on a planned release cycle and is aligned with all customers. Normally 11 platform released are performed each year. Deployment to QA is performed 1 week prior to production release.

All development changes are document in version control and attached to the release.



1.5.1 Development

This phase involves tasks in a sprint where a feature branch is created, developed, tested, and built. A pull request is then created.

1.5.2 Review

In this phase, the code undergoes a review process including completion checks. If it passes, it gets approved to move to QA; otherwise, it goes back to the development phase.

1.5.3 Validation

The code is deployed to QA where automated tests including bug check, security check, and unit tests are performed. If these tests are successful, a test report is created and approved for production.

1.5.4 Approval

This phase involves planning the deployment. Release notes are prepared, and stakeholders are notified about updates or downtime. Deployment is then approved.

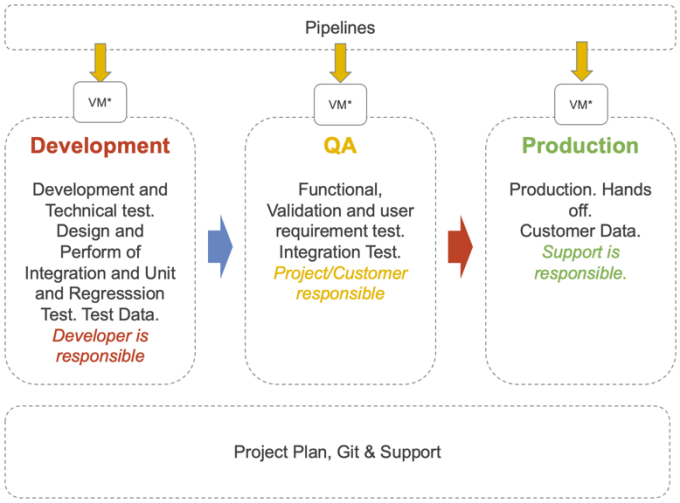
1.5.5 Production Release

The final phase where functional checks are performed after deploying to production. Production is likes QA never touched by humans and all changes are performed on a code first basis.

1.5.6 Types of Testing

Here is a brief description of who does what test and what the test includes for each environment:

- Development
  - The developer is responsible for development and technical tests. They perform design and unit integration and unit and regression tests. Test data is involved, and architect pull request approval is required. Test data is used.
- QA (Quality Assurance)
  - The project/customer is responsible for functional, validation, and user requirement tests as well as integration tests. Limited data is used.
- Production
  - DevOps teams deploys to production and validation of environment. Production is hands-off with; human access requires prior approval. Live complete dataset.



1.6. Security

1.6.1 Principle

Security is based on **NTT DATA's security practices**, with the option for external audits and ISO 27017 certification. This principle highlights the importance of adhering to NTT DATA's established security protocols and standards. It also underscores the value of optional external audits and compliance with ISO 27017 to ensure robust data security and privacy. Additional compliance checks can be ordered on demand. While the solution itself is not directly covered by ISO 27017 certification, it is built on and adheres to the principles outlined in ISO 27017. This ensures that our cloud services maintain the highest standards of security and privacy, providing our clients with confidence in the protection of their data.

Data is encrypted at rest and in transit, and customer data is isolated and placed in separate subscriptions, resource groups and objects: This concept ensures that data is encrypted both when it is stored (at rest) and when it is being transferred (in transit). Additionally, customer data is kept isolated to prevent unauthorized access or leaks.

All access is **logged and monitored**: Every attempt to access the system or data within it is recorded and monitored. This allows for real-time oversight of system interactions, enabling quick responses to any unauthorized or suspicious activity.

MFA is required for administrative access: Multi-Factor Authentication (MFA) adds an extra layer of security by requiring two or more verification methods to gain administrative access, reducing the risk of unauthorized entry.

Entra ID groups are used to manage resource access: Entra ID groups help in organizing users, managing their permissions, and controlling their access levels efficiently, ensuring that resources are accessed only by authorized personnel.

Platform Data is stored in **Azure Western Europe**: Indicates geographical location specificity for data storage; utilizing Microsoft Azure's cloud services in the Netherlands ensures compliance with regional data protection regulations. Customers can choose to store data in a region of their choice.

Principles of 4 eyes and **segregation of duties** of any administrative or developer access: Ensures that no single individual has complete control over a process or task. It requires at least two individuals to approve or complete a task which enhances accountability and reduces risks associated with fraud or errors.

Architecture is actively updated by Azure and vulnerabilities are scanned: The system architecture hosted on Azure undergoes regular updates for optimal performance while continuously scanning for vulnerabilities ensuring robust security measures are always in place.

Developers are required to fulfil security training. **Developers follow OWASP**: Developers undergo mandatory training on security protocols; adherence to Open Web Application Security Project (OWASP) guidelines ensures application safety from common threats & vulnerabilities.

Services based architecture. No physical servers.: Eliminates risks associated with physical servers like damage or theft as all services are hosted online.

Names of administrators are logged and maintained in a central location.

1.6.2 Secret Lifecyle

The main principles of secret, certificates and key lifecycle is to have very short renewal periods. This ensure that if a secret is leaked, it only has a very short valid lifetime.

1. 7 days
    - Secrets
    - Developer Access Credentials
    - Webhook IDs
    - Database Server Credentials
2. 30 days
    - Storage Access Keys
    - Service Principals
    - Database Admin Account
3. 60 days
    - SSL certificates

1.6.3 User security roles

This section describes the user roles in Momenttum Launchpad

- Momenttum Admins - Can Access all Tenants, but not all data. This role is limited to a few users in NTT DATA
- Tenant Admins - Can Access resources that belong to the tenants. This is a tenant wide setting and applies to the tenant. A user could be admin in one tenant and user in a different tenant.
- Mixed access levels only apply on Organisation, Sections and Subsection levels, mixed access is inclusive, which means the highest privilege in a hierarchy applies to the whole hierarchy.

1.7. Backup

We backup all data from core systems and perform a yearly recovery test. In customer systems we follow the same method unless a different process is agreed.

1.8. Ethics

The developers behind momenttum are working with data from many sources and to support our high standard of ethics we have enshrined a set of data ethics principles to support ethical decision-making in the use of our solutions.

Our data ethics principles cover all types of data collected, analysed, stored, shared, and otherwise processed. Our principles draw on established concepts in privacy, human rights, and business ethics to ensure we work with data in a way that maximizes benefits and minimizes harm for individuals and society. Our business solutions adopt data ethics through policies, training, communication, monitoring activities and audits.

Our ethics are grouped into 5 categories:

- Equality, fairness & non-discrimination
- Transparency
- Responsible data sharing
- Autonomy
- Dignity

In the documents “Data Ethics” and “Data Science Ethics” these topics are explained in detail.

1.9. Legal disclaimer

The information in this document is proprietary to NTT DATA Business Solutions. No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of NTT DATA Business Solutions. The information contained herein may be changed without prior notice.

NTT DATA Business Solutions assumes no responsibility for errors or omissions in this document. NTT DATA Business Solutions does not warrant the accuracy or completeness of the information, text, graphics, links, or other items contained within this material. This document is provided without a warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, or non-infringement.

NTT DATA Business Solutions shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials. This Platform functional description Date Page 5 of 5 www.nttdata-solutions.com limitation shall not apply in cases of intent or gross negligence The statutory liability for personal injury and defective products is not affected.

NTT DATA Business Solutions has no control over the information that you may access through the use of hot links contained in these materials and does not endorse your use of third-party web pages nor provide any warranty whatsoever relating to third-party web pages