ETSI TS 6756 V0.0.2 (2023-06)

MTS: Continuous Auditing Based Conformity Assessment for AI-enabled systems

<

**TECHNICAL SPECIFICATION**

Reference

DTS/MTS-6756

Keywords

Conformity,assesment

***ETSI***

650 Route des Lucioles

F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B

Association à but non lucratif enregistrée à la

Sous-préfecture de Grasse (06) N° w061004871

***Important notice***

The present document can be downloaded from:  
<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at [www.etsi.org/deliver](http://www.etsi.org/deliver).

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at <https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:  
<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

If you find a security vulnerability in the present document, please report it through our

Coordinated Vulnerability Disclosure Program:

<https://www.etsi.org/standards/coordinated-vulnerability-disclosure>

***Notice of disclaimer & limitation of liability***

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or

other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

***Copyright Notification***

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.  
The content of the PDF version shall not be modified without the written authorization of ETSI.  
The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI yyyy.

All rights reserved.

Contents

Intellectual Property Rights 4

Foreword 4

Modal verbs terminology 4

Executive summary 4

Introduction 4

1 Scope 5

2 References 5

2.1 Normative references 5

2.2 Informative references 5

3 Definition of terms, symbols and abbreviations 5

3.1 Terms 5

3.2 Symbols 6

3.3 Abbreviations 6

4 CABCA Fundamentals 6

4.1 Definition of Continuous Auditing Based Conformity Assessment (CABCA) 6

4.2 Differences of CABCA compared to traditional "point-in-time" and self-assessment methods 6

Principles 7

CABCA Assessment Process 7

Architecture, Roles, Procedures 7

Compliance 7

Risk-based quality requirement definition 8

Operationalization of quality requirements 8

Automated Assessment 8

Automated quality assessments using an assessment engine 8

Outcome of the assessments (update of conformity status) 9

CABCA in MLOps 9

Quality requirement definition based on AI-Risk-Management Frameworks 9

CABCA implementation in MLOps 9

Collection of relevant artifacts generated during the ML lifecycle 9

Evaluation and reporting of collected evidence via an API and Assessment Engine 10

Publication of a live conformity status on a third-party platform 10

Conclusion 10

Summary of the key points discussed in the document 10

Annex A (normative or informative): Title of annex 11

Annex (informative): Bibliography 12

Annex (informative): Change History 13

History 14

# Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server ([https://ipr.etsi.org](https://ipr.etsi.org/)).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

**DECT™**, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

# Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee {ETSI Technical Committee|ETSI Project|<other>} <long techbody> (<short techbody>).

# Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](https://portal.etsi.org/Services/editHelp!/Howtostart/ETSIDraftingRules.aspx) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

# Executive summary

# Introduction

# 1 Scope

The present document specifies the key aspects of Continuous Auditing Based Conformity Assessment (CABCA) as an audit methodology to evaluate and asses an organization's conformity to relevant standards and regulations.

The present document specifies:

* Principles underlying CABCA, including independence, reliability, stakeholder trust, and transparency.
* CABCA assessment process, covering architecture, roles, and procedures.
* Outcome of the assessments, including the issuance or revocation of conformity status.

# 2 References

## 2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non‑specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at <https://docbox.etsi.org/Reference>.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1]

## 2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non‑specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1]

# 3 Definition of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the [following] terms [given in ... and the following] apply:

**Conformity Assessment:** The process of evaluating and determining whether a product, service, or system complies with specified requirements, such as standards or regulations.

**Continuous Auditing:** An ongoing process of collecting, analyzing, and reporting audit-related information, typically conducted in real-time or near-real-time, to provide stakeholders with timely insights into an organization's operations and compliance status.

**MLOps:** A set of practices and methodologies for managing the lifecycle of machine learning (ML) models, including development, deployment, and maintenance.

**AI-Risk-Management Frameworks:** Guidelines and best practices for identifying, assessing, and mitigating risks associated with artificial intelligence (AI) systems, such as machine learning models, to ensure their safe and responsible use.

**Assessment Engine:** A software component that automates the evaluation of collected artifacts against predefined quality criteria to generate quality assessment outcomes.

## 3.2 Symbols

For the purposes of the present document, the [following] symbols [given in ... and the following] apply:

## 3.3 Abbreviations

For the purposes of the present document, the [following] abbreviations [given in ... and the following] apply:

**CABCA:** Continuous Auditing Based Conformity Assessment, a dynamic approach to evaluating and confirming an organization's adherence to relevant standards and regulations.

**ML:** Machine Learning, a subset of artificial intelligence that involves the development of algorithms and models capable of learning from and making predictions or decisions based on data.

**AI:** Artificial Intelligence, the development of computer systems capable of performing tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and natural language understanding.

**API:** Application Programming Interface, a set of rules and protocols that allow software components to interact and share data with each other.

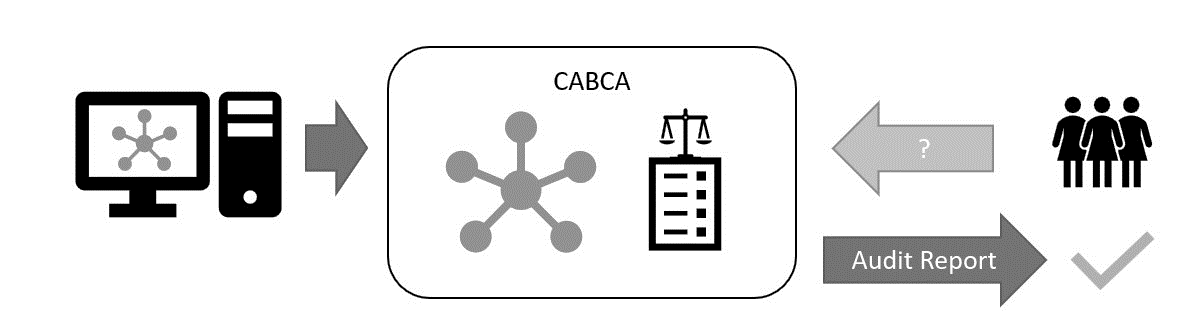
**GDPR:** General Data Protection Regulation, a comprehensive data protection law in the European Union that governs the processing and handling of personal data.

**ISO:** International Organization for Standardization, an independent, non-governmental international organization that develops and publishes international standards.

# 4 CABCA Fundamentals

## 4.1 Definition of Continuous Auditing Based Conformity Assessment (CABCA)

Continuous Auditing Based Conformity Assessment (CABCA) is a dynamic approach to evaluating and confirming an organization's adherence to relevant standards, rules, and regulations, including those set by bodies like ISO and ETSI, sector-specific regulations, market rules, and customer demands. CABCA is designed to provide uninterrupted assessment and surveillance of an organization's systems and operations, thereby establishing an ongoing compliance status and cultivating trust among stakeholders. This methodology is particularly crucial for systems with fast evolution cycles, such as Machine Learning (ML) systems, where traditional semi-annual certifications fail to keep pace with the rapid updates and changes.



## 4.2 Principles of CABCA

At the heart of CABCA are its guiding principles, which serve as the foundation for the methodology's effectiveness and reliability. These principles can be boiled down to three main pillars: Ongoing Conformity, Stakeholder Trust, and Adaptability.

Ongoing Conformity

Traditional conformity assessment methodologies often offer a 'snapshot' of compliance at a particular moment in time. While this approach may suffice for static or slowly evolving systems, it's inadequate for rapidly changing environments, particularly in sectors like Machine Learning. CABCA shifts this paradigm by focusing on a real-time, continuous understanding of an organization's adherence to standards. This principle aligns closely with the core process of Continuous Assessment, which leverages real-time data for an uninterrupted compliance status. This approach not only provides a more nuanced and current view of compliance but also integrates seamlessly with risk-based quality requirements and risk management strategies.

Stakeholder Trust

Trust is a crucial element for any organization's success, especially in today's rapidly changing landscape. The CABCA methodology places a high value on transparency and open communication with stakeholders, which in turn fosters trust. Consistent reporting and auditing through the Audit Report Publication process keep stakeholders updated, enhancing awareness and understanding of the organization's compliance activities and status. This continuous flow of reliable information boosts confidence among stakeholders, allowing them to make well-informed decisions. This principle also incorporates the benefits of greater focus on stakeholder communication, ensuring that all parties involved have a clear understanding of the organization's commitment to maintaining compliance with relevant standards.

Adaptability

With the fast-paced nature of technological advancement, any auditing system must be adaptable to remain relevant. CABCA is designed to be flexible enough to keep up with changes in both the internal systems being audited and the external regulatory landscape. This flexibility is especially relevant for dynamic and probabilistic systems like Machine Learning, where static auditing methods fall short. The Initialization process in CABCA allows organizations to define and, if necessary, redefine the scope of the audit based on evolving regulations, market demands, or technological advancements. This adaptability emphasizes continuous improvement, enabling organizations to refine their compliance processes in response to emerging risks and changes in the regulatory landscape.

In summary, these core principles—Ongoing Conformity, Stakeholder Trust, and Adaptability—not only guide the CABCA methodology but also offer distinct advantages over traditional "point-in-time" and self-assessment methods. They ensure a more standardized yet flexible approach to conformity assessment, instill confidence among stakeholders, and provide an infrastructure for organizations to continually improve and adapt to new compliance challenges. These principles align with the overarching goals of CABCA, including its focus on maintaining transparency and demonstrating an ongoing commitment to quality and compliance.

## Description of the CABCA Process

The CABCA is a dynamic methodology designed to keep pace with the rapid updates typical of Machine Learning (ML) systems. Unlike traditional compliance mechanisms like semi-annual certifications, CABCA is agile, allowing organizations to stay compliant with legislative requirements, ETSI standards, or self-imposed quality.

At its core, CABCA revolves around the concept of 'Operationalization,' which involves translating high-level compliance standards into machine-readable metrics. This process kicks off by identifying 'Quality Dimensions' such as fairness, data protection, or reliability that the ML system must adhere to. Associated risks within these dimensions are then identified.

To mitigate these risks, specific 'Requirements' are set for the AI system. Corresponding 'Metrics' are defined to quantitatively assess the fulfillment of these requirements. Tools are often employed to carry out 'Measurements,' generating values for these metrics.

Automated assessments then evaluate these measurements to see if the requirements have been met and consequently if the risks have been mitigated. Each assessment iteration provides a 'Conformity Status,' available for stakeholders like customers, legislators, or internal teams to review.

CABCA builds trust by being transparent and effectively communicating this continuously updated status. It also accommodates third-party audits for an unbiased validation of the compliance status. Overall, CABCA ensures that organizations not only achieve but maintain compliance in a landscape where technological and regulatory conditions are in flux.

## Modes of CABCA

* Self-Assessment: In this mode, the organization itself acts as both the auditee and the auditing party to conduct internal audits for ongoing compliance with legislative requirements and quality goals.
* Third-Party Audit: Here, an external accredited body serves as the auditing party to validate the organization's adherence to standards.
* Certification: In this mode, CABCA assists organizations in procuring certifications from accredited bodies, ensuring a comprehensive approach to compliance.

## Architecture, Roles, Procedures

* Auditee: Responsible for defining the scope of CABCA, which reflects the quality requirements for the ML/AI system. Implements technical measurements to provide evidence to the auditor.
* Auditing Party: Conducts the audit under the rules set by the Conformity Specification Publisher. Verifies the scope provided by the auditee and facilitates automated measurements and assessments.
* Conformity Status Publishing Entity: Provides stakeholders with the current conformity status of the ML/AI system. This entity could be the same as the auditing party or a separate entity depending on the mode of CABCA.

# Operationalization in CABCA

## Capabilities of Conformity Specifications for CABCA

Conformity Specifications refer to high-level documents that specify the types of conformity required in a particular context. These could be international standards like those set by ISO or ETSI, national standards, or other authoritative sources that outline the conditions or guidelines that must be met. The sources of these specifications can vary widely, depending on factors such as the industry, market standatds, and specific organizational needs. Examples of sources include:

* ISO, ETSI, etc. Standards: High-level international standards that often serve as a framework for conformity in various sectors.
* National Standards: Developed by national organizations like NIST in the United States.
* Industry Guidelines: Established best practices within specific industries.
* Internal Policies: Guidelines set within an organization.
* Legislative Requirements: Federal, state, or local laws that mandate certain types of conformity.
* Quality Assurance Protocols: These could be part of internal governance or subject to external audits.
* Product Certifications: Such as UL, CE, or Commo criteria labels that certify a product meets certain standards.
* Military Standards (MIL-STD): These are relevant to military applications.
* Healthcare Regulations: Such as HIPAA in the United States.
* Financial Regulations: Sarbanes-Oxley for corporate governance, GDPR for data protection, and so on.
* Vendor Agreements: Requirements from vendors that organizations must adhere to.
* Ethical Guidelines: Ethical principles and frameworks, like fairness, transparency, and accountability, that apply specifically to AI development and deployment.
* Third-Party Audits: These may also dictate conformity requirements.
* Consumer Protection Laws: Regulations to ensure AI-driven consumer products meet safety and reliability standards, as well as disclosure requirements around data collection and usage.

Given the dynamic nature of Continuous Auditing Based Conformity Assessment (CABCA), it's imperative for these Conformity Specifications to meet specific minimal capabilities to be effective within this real-time, continuous framework.

Capabilities of Conformity Specifications for CABCA:

* Operationalizability: The specification should be able to be translated into practical steps for risk management, quality assurance, or other operational activities.
* Coverage of Key Areas: At a minimum, the specifications should cover the essential aspects relevant to the system or industry under CABCA audit.
* Credibility: The specification should have some degree of recognition or authority within its respective industry.
* Clear Licensing and Usage Policies: There must be explicit terms defining how the specification can be used.
* Capability to Provide Metrics: The specification should either directly provide measurable parameters or be interpretable in a way that such metrics can be derived.
* Auditability: The specification should contain elements that are auditable on a continuous basis, aligning with CABCA’s methodology.

The evolving nature of industries like machine learning demands a flexible approach to conformity assessment. The listed minimal capabilities are essential for integrating a conformity specification into the CABCA framework.

## Process of Operationalization

Operationalization in this context refers to the translation of Conformity Specifications into actionable steps and metrics, aligning with the capabilities outlined above. This ensures that the ML system remains compliant with set standards and quality goals, thereby fostering trust among stakeholders.

**Align with Conformity Specifications: The first step is to align the operationalization process with the capabilities of the chosen Conformity Specifications. This ensures that the specifications are not only credible and comprehensive but also operationalizable.**

**Identify Compliance Dimensions:** Based on the Conformity Specifications, identify the dimensions that the ML system needs to be compliant with. These could be the 'Quality Dimensions' such as fairness, data protection, and reliability, as previously described in the CABCA process.

**Risk Identification:** Utilize the Conformity Specifications to recognize the risks associated with each dimension. This step is crucial for setting the stage for risk mitigation strategies.

**Define Quality Requirements:** Translate the Conformity Specifications into specific 'Requirements' for the ML system. These requirements serve as the criteria against which the system’s compliance will be assessed and should be aligned with the capabilities of the specifications.

**Define Metrics and Measurements:** For each quality requirement, define corresponding 'Metrics' and assign 'Measurements' to them. These metrics and measurements should be derived from or aligned with the Conformity Specifications, ensuring they are both auditable and applicable.

The operationalization process starts with the selection of relevant 'Quality Dimensions' from the conformity specification such as fairness, data protection, and reliability. These dimensions serve as high-level categories for auditing the ML system. With the Quality Dimensions identified, the next phase involves recognizing the risks associated with each dimension. This is a crucial part of the process as it informs what will be closely monitored during the continuous auditing. Using the Conformity Specifications as a reference, risks are identified and appropriate mitigation strategies are outlined. These strategies then form the basis for the specific 'Requirements' that the ML system needs to meet. In other words, each risk mitigation strategy is translated into an actionable requirement. Once requirements are clearly defined, metrics are then established for each. These metrics provide measurable parameters that reflect how well the system is adhering to each requirement. After defining the metrics, the subsequent task is to assign methods for measuring these metrics. Measurements are determined for each metric, specifying the techniques, tools, or procedures to be used in the ongoing assessment.



### Capabilities for Dimensions

* Granularity: Ability to dissect dimensions into smaller, manageable components for easier auditing.
* Adaptability: Flexibility to adapt dimensions to evolving technology or business scenarios.

### Capabilities for Risks

* Mitigatability: Ability to implement risk mitigation strategies.
* Risk Traceability: Ability to trace each mitigation action back to its originating risk and source specification.

### Capabilities for Requirements

* Clarity: Clear and unambiguous definition of requirements.
* Measurability: Presence of associated metrics for each requirement.
* Requirement Traceability: Ability to trace each requirement back to its source specification.
* Frequency: Specification of how often the conformity of a requirement needs to be evaluated.

### Capabilities for Metrics

* Precision: Clear and unambiguous definition of metrics.
* Relevance: Direct relation to the requirements and dimensions being measured.
* Unit Specification: Definition of a unit for the metric's value.
* Measurability: Capability for the metric to be measured.

### Capabilities for Measurements

* Automation: High degree of automated tool-based measurement.
* Output Unit Consistency: Output unit should match the unit specified by the metric.
* Automated Result Collection: Automated collection of measurement results for continuous auditing.
* Evidence Retention: Specification of the time span for retaining measurement results as evidence.

## Documentation of the Operationalization

Documentation serves two primary purposes in the process of CABCA. Firstly, it is designed to communicate the operationalization and the selections and decisions made clearly to stakeholders, providing them with an informed understanding of ongoing conformity measures. Secondly, the documentation acts as a configuration file, which later assists in automated assessment procedures.

### Capabilities of Documentation

* Transparency: Clearly outlines the process of translating Conformity Specifications into operational steps, thus eliminating ambiguity.
* Traceability: Allows for each metric, requirement, and decision to be traced back to the original Conformity Specifications and key dimensions.
  + Clearly articulate the operationalization methods used.
  + Reduce regulatory uncertainties by offering precise interpretations of Conformity Specifications.
  + Clearly state quality requirements and expectations.
  + Clearly define the scope of each assessment.
* Configurability: Enables straightforward translation into automated assessments by acting as a configuration file.

### Structure

The documentation is organized in a tree structure:

* Root: The root of the tree is the Conformity Specification, serving as the basis for all ensuing steps.
  + Node 1: Dimensions
  + Comment: Description and justification for the choice of each dimension.
  + Confidentiality Flag: Indicates if the information is confidential.
  + Node 1.1: Requirements
    - Comment: Details of how dimensions translate into specific requirements.
    - Confidentiality Flag: Indicates if the information is confidential.
    - Frequency of assesment
    - Node 1.1.1: Metrics
      * Comment: Explanation of how metrics quantify the requirements.
      * Confidentiality Flag: Indicates if the information is confidential.
      * Node 1.1.1.1: Measurements
        + Comment: Methods for taking measurements for the metrics.
        + Confidentiality Flag: Indicates if the information is confidential.

Yaml Representation:

Root:

ConformitySpecification:

Comment: "Description and justification for Conformity Specification"

ConfidentialityFlag: false

Dimensions:

Node1:

Comment: "Description and justification for the choice of each dimension"

ConfidentialityFlag: false

Requirements:

Node1\_1:

Comment: "Details of how dimensions translate into specific requirements"

ConfidentialityFlag: false

Frequency: -1

Metrics:

Node1\_1\_1:

Comment: "Explanation of how metrics quantify the requirements"

ConfidentialityFlag: false

Measurements:

Node1\_1\_1\_1:

Comment: "Methods for taking measurements for the metrics"

ConfidentialityFlag: false

JSON Representation:

{

"Root": {

"ConformitySpecification": {

"Comment": "Description and justification for Conformity Specification",

"ConfidentialityFlag": false,

"Dimensions": {

"Node1": {

"Comment": "Description and justification for the choice of each dimension",

"ConfidentialityFlag": false,

"Requirements": {

"Node1\_1": {

"Comment": "Details of how dimensions translate into specific requirements",

"ConfidentialityFlag": false,

"Metrics": {

"Node1\_1\_1": {

"Comment": "Explanation of how metrics quantify the requirements",

"ConfidentialityFlag": false,

"Measurements": {

"Node1\_1\_1\_1": {

"Comment": "Methods for taking measurements for the metrics",

"ConfidentialityFlag": false

}

}

}

}

}

}

}

}

}

}

}

# CABCA Assessment Process

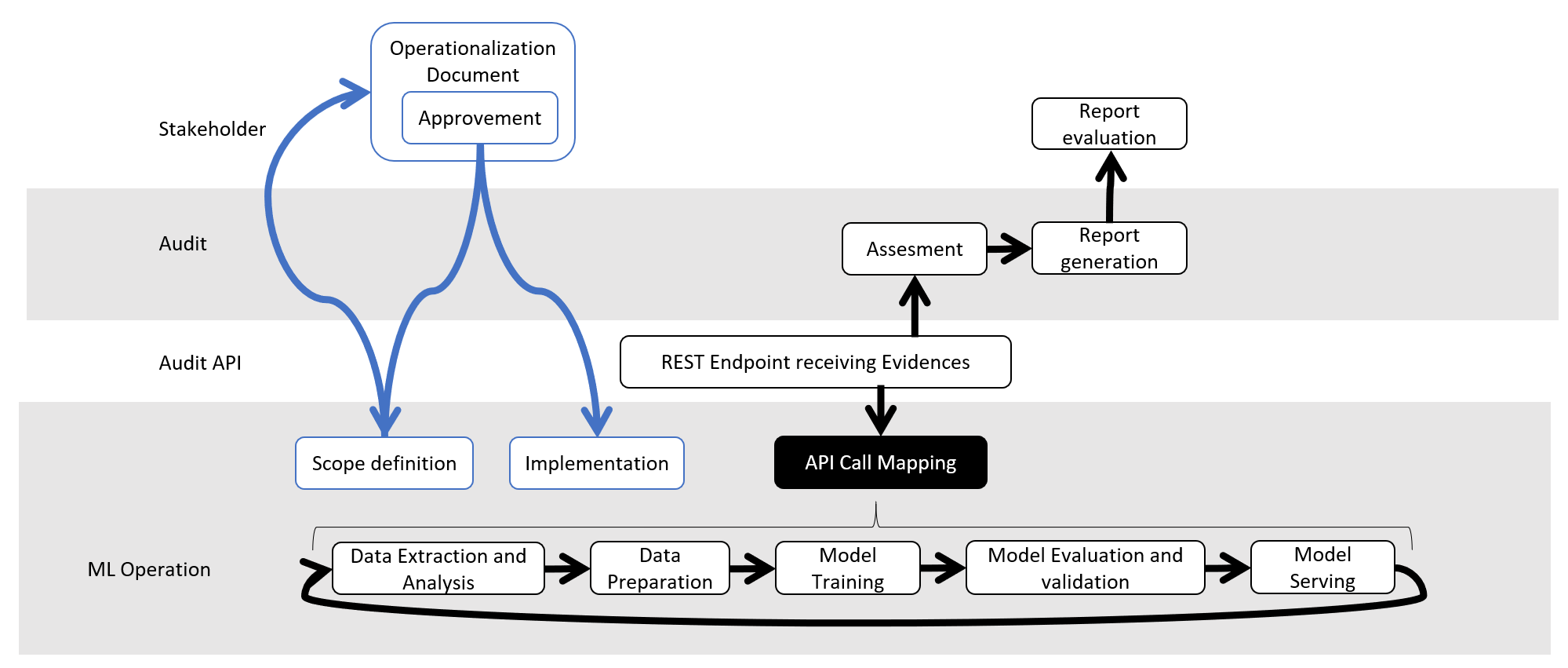
The CABCA is designed to provide an ongoing auditing of systems to ensure they meet compliance standards, quality goals, and manage risks effectively. This chapter elucidates the steps involved in CABCA assessment, focusing on how the Operationalization Documentation serves as a configuration file for the assessment process.

Initial Setup: Using Operationalization Documentation as Configuration File

Before initiating the assessment process, the Operationalization Documentation, which lays down dimensions, requirements, metrics, and measurements, is ingested into the CABCA assessment system as a configuration file. This enables automatic interpretation of audit criteria, metrics, and the corresponding measurements.

Frequency of Assessments

The frequency for each assessment varies based on individual needs and is attached to each requirement in the configuration file. The assessment process is invoked every time a new model is deployed. During the operation phase, the frequencies for assessment are explicitly defined within the Operationalization Documentation.



## The Assessment Process

1. Artifacts Production and Usage: Artifacts such as log files, model weights, and data samples are produced or utilized at various phases in the ML life cycle, serving as inputs for the measurements. Some artifacts yield measurement results directly through parsing, while others might necessitate comprehensive test suites to extract accurate data.
2. Measurement Using Artifacts as Inputs: Measurements are performed using artifacts. The results of these measurements are then prepared for transmission to the auditing entity. This step ensures real-time or frequency-based evaluations, triggered either after specific cycles such as model deployment or based on predefined intervals tied to individual quality requirements.
3. Mapping Measurement Results to ML Quality Assessment API: Auditors automatically evaluate the received measurement data, utilizing the Operationalization Documentation as a configuration file. The evaluation is against predefined metric thresholds, which reflect the quality goals of the ML system. This automated process ensures that the evaluation is not only fast but also consistent with the quality goals outlined in the Conformity Specifications.
4. Evaluation of Evidence: Auditors evaluate the received evidence and compare it against predefined metric thresholds, which reflect the system's quality goals. This evaluation process is automated, making use of the configuration file to understand what constitutes compliance or deviation.
5. Report Generation: After the evaluation, a detailed report is generated. This report outlines each measurement result, mapping it back to the quality dimensions and requirements specified in the Operationalization Documentation. It concludes with a verdict on the system's compliance status relative to its quality goals.

## ML Quality Assessment API

The Machine Learning (ML) Quality Assessment API serves as a RESTful interface, engineered to standardize and unify the reporting of disparate quality measurements across various ML systems. This API facilitates an integrative communication, bridging the gap between the measuring entity and the ML quality assessment entity. It fosters a uniform understanding and exchange of diverse measurement data, converging them onto the same conceptual plane, which ultimately enhances the holistic evaluation and refinement of ML systems.

The API consists of the following endpoints:

### /models

Endpoints related to information about the machine learning models:

* /models/architecture: Endpoint for model architecture information.
* /models/weights: Endpoint for model weight information.
* /models/metadata: Endpoint for model metadata information.
* /models/evaluation: Endpoints related to the evaluation of the machine learning models:
  + /models/evaluation/performance\_metrics: Endpoint for performance metrics of the model.
    - /models/evaluation/performance\_metrics/accuracy: Endpoint for accuracy information of the model. (Classification Metrics, Image Recognition Metrics, Natural Language Processing Metrics, Common Metrics)
    - /models/evaluation/performance\_metrics/precision: Endpoint for precision information of the model. (Classification Metrics, Image Recognition Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Metrics)
    - /models/evaluation/performance\_metrics/recall: Endpoint for recall information of the model. (Classification Metrics, Image Recognition Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Metrics)
    - /models/evaluation/performance\_metrics/f1-score: Endpoint for F1-score information of the model. (Classification Metrics, Image Recognition Metrics, Natural Language Processing Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Metrics)
    - /models/evaluation/performance\_metrics/mean\_absolute\_error: Endpoint for Mean Absolute Error information of the model. (Anomaly Detection Metrics, Regression Metrics, Common Metrics)
    - /models/evaluation/performance\_metrics/root\_mean\_squared\_error: Endpoint for Root Mean Squared Error information of the model. (Regression Metrics, Image Recognition Metrics, Common Metrics)
    - /models/evaluation/performance\_metrics/precision\_recall\_curve: Endpoint for Precision-Recall Curve information of the model. (Anomaly Detection Metrics, Recommender System Metrics, Common Metrics)
  + /models/evaluation/error\_rates: Endpoint for error rates of the model.
  + /models/evaluation/robustness: Endpoint for information on the robustness of the ML model, including its ability to handle unexpected input, changes in the input data distribution, and unexpected events in real-world scenarios.
  + /models/evaluation/explainability: Endpoints related to explainability of the machine learning models:
    - /models/evaluation/explainability/feature\_contributions: Endpoint for feature contribution information of the model.
    - /models/evaluation/explainability/permutation\_importance: Endpoint for permutation importance information of the model.
    - /models/evaluation/explainability/feature\_importance: Endpoint for feature importance information of the model. (Classification Metrics, Regression Metrics, Clustering Metrics, Natural Language Processing Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Methods)
    - /models/evaluation/explainability/partial\_dependence\_plots: Endpoint for partial dependence plot information of the model. (Classification Metrics, Regression Metrics, Clustering Metrics, Natural Language Processing Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Methods)
    - /models/evaluation/explainability/shap\_values: Endpoint for SHAP value information of the model. (Classification Metrics, Regression Metrics, Clustering Metrics, Natural Language Processing Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Methods)
    - /models/evaluation/explainability/lime\_explanations: Endpoint for LIME explanation information of the model. (Classification Metrics, Regression Metrics, Clustering Metrics, Natural Language Processing Metrics, Image Recognition Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Methods)
    - /models/evaluation/explainability/global\_explanations: Endpoint for global explanation information of the model. (Classification Metrics, Regression Metrics, Clustering Metrics, Natural Language Processing Metrics, Image Recognition Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Methods)
    - /models/evaluation/explainability/local\_explanations: Endpoint for local explanation information of the model. (Classification Metrics, Regression Metrics, Clustering Metrics, Natural Language Processing Metrics, Image Recognition Metrics, Anomaly Detection Metrics, Recommender System Metrics, Common Methods)

### /data

Endpoints related to information about the data used in the machine learning models:

* /data/source: Endpoint for data source information.
* /data/quality: Endpoint for data quality information.
* /data/distribution: Endpoint for data distribution information.
* /data/preprocessing: Endpoint for data preprocessing information.
* /data/evaluation: Endpoints related to the evaluation of the data used in the machine learning models:
  + /data/evaluation/syntactic\_accuracy: Endpoint for syntactic accuracy information of the data.
  + /data/evaluation/semantic\_accuracy: Endpoint for semantic accuracy information of the data.
  + /data/evaluation/completeness: Endpoint for data completeness information.
  + /data/evaluation/consistency: Endpoint for data consistency information.
  + /data/evaluation/integrity: Endpoint for data integrity information.
  + /data/evaluation/uniqueness: Endpoint for data uniqueness information.
  + /data/evaluation/data\_distribution: Endpoint for data distribution information.
  + /data/evaluation/data\_balance: Endpoint for data balance information.
  + /data/evaluation/anomaly\_detection: Endpoint for information on anomaly detection in the data.
  + /data/evaluation/outlier\_detection: Endpoint for information on outlier detection in the data.

#### /training

Endpoints related to information about the training process:

* /training/algorithm: Endpoint for training algorithm information.
* /training/hyperparameters: Endpoint for hyperparameter information used in the training process.
* /training/performance\_metrics: Endpoint for performance metrics of the training process.

#### /inference

Endpoints related to information about the inference process:

* /inference/algorithm: Endpoint for inference algorithm information.
* /inference/runtime\_performance\_metrics: Endpoint for runtime performance metrics of the inference process.
* /inference/real\_time\_predictions: Endpoint for information on real-time predictions of the model.

## Persistence of Assessment Results

The results of each CABCA assessment need to be stored persistently for future reference, quality tracking, and for proving compliance in audits. The scope of this persistence includes:

* Measurement Results: All metrics and their corresponding values are stored.
* Evaluation Outcomes: The outcomes of the comparisons between metrics and pre-defined quality goals are archived.
* Assessment Reports: The final reports generated after each assessment are saved.

The retention policies for these artifacts are defined according to industry standards, legal requirements, and organizational policies.

### Capabilities for Persistence

* Searchability: All stored data should be easily searchable.
* Security: Ensure that the stored data is encrypted and accessible only by authorized personnel.
* Traceability: Each stored item should be traceable back to the specific assessment cycle and corresponding Operationalization Documentation.
* Automated Archiving: An automated archiving system should be in place to manage the lifecycle of stored data.

## Building Stakeholder Trust through CABCA

CABCA builds trust by fostering transparency, facilitating effective communication, regularly updating compliance status, and incorporating third-party audits.

### Transparency Requirements

* Communicate risk mitigation methods.
* Reduce regulatory uncertainties by providing interpretations.
* Communicate quality requirements.
* Be transparent about how conformity is implemented and assessed, the scope of the assessment, and updates in the implementation.
* Openly communicate the results of the Conformity Assessment without revealing crucial insights into the ML System.

### Communication Requirements

* conformity status or other pertinent information.
* Define communication channels and the scope for each channel.
* Define the frequency of communication updates and their granularity.

### Inclusion of Third-Party Expertise

* Involve third-party audits for unbiased assessment and verification of the system’s compliance.
* Define the degree of third-party involvement in conformity implementation, assessment, and setup.
* Ensure separation between the ML System and the Audit System.

# Automated Assessment

Collection of relevant artifacts

This section explains how relevant artifacts, such as data, models, and performance metrics, are collected and stored in CABCA.

* Explanation of relevant artifacts used in the CABCA assesment process
* Methods for collecting artifacts, including automating the monitoring and collection of data from different parts of the system
* Overview of secure protocols and access controls used to ensure data security and privacy during data transfer and storage
* Explanation of how the collected artifacts are used in the automated quality assessments.

## Automated quality assessments using an assessment engine

This section describes how the assessment engine uses pre-defined quality criteria to perform automated quality assessments.

* Overview of the automated quality assessment process using an assessment engine
* The role of the assessment engine in conducting continuous quality assessments
* Explanation of the pre-defined quality criteria used by the assessment engine
* Continuous quality assessments based on pre-defined quality requirements
* Implementing measurements for quality
* Overview of the data inputs required for the assessment engine (relevant artifacts)
* Mapping results of multiple measurement tools and other data to a combined input for the assessment
* Discussion of the assessment engine's evaluation methodology and decision-making process
* Explanation of the outcomes of the assessment

## Outcome of the assessments (update of conformity status)

This section describes the outcome of the quality assessments, including the update of the conformity status.

* Explanation of the outcome of the assessments performed by the assessment engine
* Evaluation of measurement results, taking artifacts as input, against predefined values derived from expert knowledge and risk assessment
* Issuance of conformity status if the measurement results align with predefined values, indicating compliance with required standards
* Revocation or adjustment of conformity status if the measurement results do not meet predefined values, highlighting areas for improvement and increased risk management
* Discussion of the process for updating or revising the pre-defined quality requirements and the corresponding impact on the assessment outcome

# CABCA in MLOps

## Quality requirement definition based on AI-Risk-Management Frameworks

This section explains how quality requirements for ML systems are defined in CABCA using AI-Risk-Management frameworks.

* Overview of AI-Risk-Management Frameworks and their relevance for CABCA
* Definition of quality requirements for CABCA in the context of AI-Risk-Management
* Integration of AI-Risk-Management Frameworks with CABCA processes
* Alignment of CABCA quality requirements with AI-Risk-Management best practices and standards
* Definition of risk-based quality requirements for AI systems, models and their components
* Continuous updating and refinement of quality requirements based on changes in the AI-Risk-Management landscape.

## CABCA implementation in MLOps

This section describes how CABCA is implemented in the context of MLOps, including the processes, procedures, and tools involved.

* Explanation of the integration of CABCA into the MLOps pipeline
* Overview of how CABCA fits into the ML lifecycle
* Discussion of the benefits of incorporating CABCA into MLOps
* Explanation of how CABCA can improve the quality and reliability of ML models
* Discussion of the role of CABCA in managing AI-related risks in MLOps
* Overview of the process for defining and implementing quality requirements for ML models in the context of CABCA.

## Collection of relevant artifacts generated during the ML lifecycle

This section explains how relevant artifacts generated during the ML lifecycle are collected and used for quality assessment in CABCA. Artifacts used for quality assessment (data, model, architecture, configurations, hyperparameters, algorithm, metrics, logs, etc.) - This subsection provides a list of the artifacts that are used for quality assessment in CABCA.

* Overview of relevant artifacts generated during the ML lifecycle
* Collection of artifacts relevant to the quality assessment process
* Examples of artifacts: data, model, architecture, configurations, hyperparameters, algorithm, metrics, logs, etc.
* Procedures and tools used to collect the artifacts
* Automated collection of artifacts at various stages of the ML lifecycle
* Record-keeping of collected artifacts for future reference.

## Evaluation and reporting of collected evidence via an API and Assessment Engine

This section describes how the collected evidence is evaluated and reported in CABCA, using an API and assessment engine.

* Overview of the process for collecting evidence related to the ML lifecycle
* Explanation of how the evidence is evaluated and reported through an API and assessment engine
* Details on how the API and assessment engine interface to produce a comprehensive evaluation report
* Discussion of how the report influences the conformity status update.

## Publication of a live conformity status on a third-party platform

This section explains how the results of the quality assessments are published in CABCA, including the publication of the conformity status.

* Explanation of the concept of publishing a live conformity status on a third-party platform
* Steps involved in publishing a live conformity on a governing body platform
* Details of the publishing platform, including its purpose, features, and accessibility
* Benefits of publishing a live status conformity on a governing body platform, such as increased transparency and credibility
* Explanation of how a live conformity status can be updated based on continuous quality assessments

# Conclusion

## Summary of the key points discussed in the document

This section provides a summary of the key points discussed in the document, highlighting the main benefits and drawbacks of CABCA.

Annex A (normative or informative):  
Title of annex

Annex (informative):  
Bibliography

Annex (informative):  
Change History

| Date | Version | Information about changes |
| --- | --- | --- |
| <Month year> | <#> | <Changes made are listed in this cell> |
|  |  |  |
|  |  |  |
|  |  |  |

# History

|  |  |  |
| --- | --- | --- |
| **Document history** | | |
| <Version> | <Date> | <Milestone> |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

*Latest changes made on 2022-03-14*