D5.12 – i4Q Manufacturing Line Data Certification Procedure v2

WP5 – BUILD: Rapid Manufacturing Line Qualification and Reconfiguration

i4Q has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958205.

Document Information

GRANT AGREEMENT NUMBER	958205	ACRONYM	i4Q	
FULL TITLE	Industrial Data Services for Quality Control in Smart Manufacturing			
START DATE	01-01-2021	DURATION	36 months	
PROJECT URL	https://www.i4q-project.eu/			
DELIVERABLE	D5.12 – i4Q Manufacturing Line Data Certification Procedure v2			
WORK PACKAGE	WP5 – BUILD: Rapid Manufacturing Line Qualification and Reconfiguration			
DATE OF DELIVERY	CONTRACTUAL	31-Dec-2022	ACTUAL	30-Dec-2022
NATURE	Other	DISSEMINATION LEVEL	Public	
LEAD BENEFICIARY	10-TUB			
RESPONSIBLE AUTHOR	Anna M. Nowak-Meitinger, Stefan Wellsandt, Robert Trevino			
CONTRIBUTIONS FROM	10-TUB, 8-BIBA, 7-IKER			
TARGET AUDIENCE	1) i4Q Project partners; 2) industrial community; 3) other H2020 funded projects; 4) scientific community			
CONTEXT/ DEPENDENCIES	This document presents the second version of D5.6 i4Q Manufacturing Line Data Certification Procedure. Its relationship to other documents is as follows: - D5.6 i4Q Manufacturing Line Data Certification Procedure: first version of i4Q solution i4Q^{LCP}. - D3.1 i4Q Data Quality Guidelines: mentions this deliverable as a measure within the activity framework. - D3.9 i4Q Data Quality Guidelines v2: mentions this deliverable as a measure within the activity framework.			
EXTERNAL ANNEXES/ SUPPORTING DOCUMENTS	None			
READING NOTES	None			

ABSTRACT	Deliverable D5.12 contains the second version of the i4Q Manufacturing Line Certification Procedure (i4Q^{LCP}), a guideline that includes an audit procedure and an IT assistance tool. The guide aims to ensure high data quality in manufacturing processes. It is developed as part of the i4Q project and includes recommendations for process reconfiguration and audit strategies. To develop the framework, the basics of audit standard ISO 19011:2018 are covered, including vocabulary, principles, and roles. In addition, prerequisites are defined and data quality standards (e.g., ISO/IEC 25012:2008, ISO/IEC 25024:2015) are used to develop audit criteria and expand the framework. The digital tool is presented as a chatbot to support audit planning, implementation, and conduct. This solution can also serve as a basis to complement existing quality certifications by supporting standardization in the field of data quality certification.

Document History

VERSION	ISSUE DATE	STAGE	DESCRIPTION	CONTRIBUTOR
0.1-0.9	07-Nov-2022	ToC	ToC Draft and first input	TUB
0.10-0.14	21-Nov-2022	Working Version	Input to all sections	TUB, BIBA
0.15-0.17	24-Nov-2022	Working Version	Formatting & Review	TUB, BIBA, IKER
1.0	25-Nov-2022	1st Draft	Available for Internal Review	TUB
1.1	30-Nov-2022	Internal Review	Internal Review	DIN, LIF
1.2	02-Dec-2022	2nd Draft	Addressing comments from internal reviewers	TUB, BIBA
1.3	09-Dec-2022	Final Draft	Submission for final draft for quality check	TUB
1.4	30-Dec-2022	Final Document	Final quality check and issue of final document	CERTH

Disclaimer

Copyright message

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS/ACRONYMS

AI	Artificial Intelligence
CEN	European Committee for Standardisation
CNC	Computer Numerical Control
CWA	CEN Workshop Agreement
D1.9	i4Q Deliverable D1.9 – Requirements Analysis and Functional Specification v2
D2.3	i4Q Deliverable D2.3 – Report on Business Viewpoint
D2.4	i4Q Deliverable D2.4 – Report on Usage Viewpoint
D3.1	i4Q Deliverable D3.1 – i4Q Data Quality Guidelines
D3.2	i4Q Deliverable D3.2 – i4Q QualiExplore for Data Quality Factor Knowledge
DB	DataBase
DBMS	DataBase Management System
DIM	Product Dimension
DIN	German Institute for Standardisation
DLC	Data Life Cycle
FSD	Function Structure Diagram
GDPR	General Data Protection Regulation
i4Q^{LCP}	i4Q Manufacturing Line Certification Procedure
i4Q^{LRT}	i4Q Manufacturing Line Reconfiguration Toolkit
i4Q^{PA}	i4Q Prescriptive Analysis Tools
i4Q^{PQ}	i4Q Data-driven Continuous Process Qualification
i4Q^{QD}	i4Q Rapid Quality Diagnosis
i4Q^{QE}	i4Q QualiExplore
IP	Intellectual Property
ISO	International Standardisation Organisation
IT	Information Technology
KPI	Key Performance Indicator
MES	Manufacturing Execution System
MS forms	Microsoft forms

PDCA	Plan-Do-Check-Act
PDF	Portable Document Format
PLC	Programme Logic Controllers
RDMS	Relational Database Management System
RPM	Revolutions Per Minute
SoA	State-of-the-art
SQL	Structured Query Language
UI	User Interface
URL	Uniform Resource Locator
QMS	Quality Management System
WP	Work package

Executive summary

Deliverable D5.12 presents the second version of i4Q Line Certification Procedure (i4Q^{LCP}). This solution provides an audit guideline that is applied to manufacturing resources (machine, cell, or manufacturing line) to measure and assess data quality in manufacturing. The goal is to ensure that all data resulting from the manufacturing processes are accurate and reliable. Therefore, recommendations for process reconfiguration, audit strategies, certificates and regulations are included.

The first three sections serve as an introduction and foundation of the guideline. The objectives and context of i4Q^{LCP} are presented and as a basis for the guideline development, standards to design i4Q^{LCP} Solution such as *ISO 19011:2018 Guidelines for auditing management systems* and the series of standards *ISO 8000 Data Quality* and *ISO 25000 Systems and software engineering* are introduced. Relevant theoretical foundations of these standards and their contribution to i4Q^{LCP} are presented. Furthermore, basic prerequisites for the use of i4Q^{LCP} guideline are described, including the certification according to quality management standard ISO 9001:2015 and the realization of capable measurement systems in industrial manufacturing as well as capable IT-infrastructures.

Section 4 presents the content of i4Q^{LCP} guideline. Therefore, ISO 19011:2018 is used to define the frame, objectives, risks, and roles as well as the planning, implementation, controlling and improvement of the audit procedure. Furthermore, the standards ISO/IEC 25012:2008, ISO/IEC 25024:2015 are used to define the audit criteria. Data quality models and characteristics are adapted, combined, and presented to structure and develop the audit criteria and cluster them into sets of requirements for each phase of the data life cycle (DLC).

Section 5 presents the second part of the solution, provided as an IT-based assistant Audit Advisor (chatbot) that guides the user through the process of assessing compliance with the audit criteria. This tool includes the list of audit criteria questions and an automatic evaluation of the level of conformity.

The content structure of this guideline is shown in **Figure 1**. The activities in the five main sections, Introduction, Standards for i4Q^{LCP} guideline development, Prerequisites for use of i4Q^{LCP}, i4Q^{LCP} *Manufacturing Line Data Certification Procedure* and *Digital support tools* may be conducted sequentially or in parallel. Sections with same colouring describe connections of the established standards to the use in the i4Q^{LCP} guideline.

This document i4Q D5.12 v2 is an update of v1 of D5.6., for this reason it contains information of the 1st version together with the updates developed in this 2nd version.

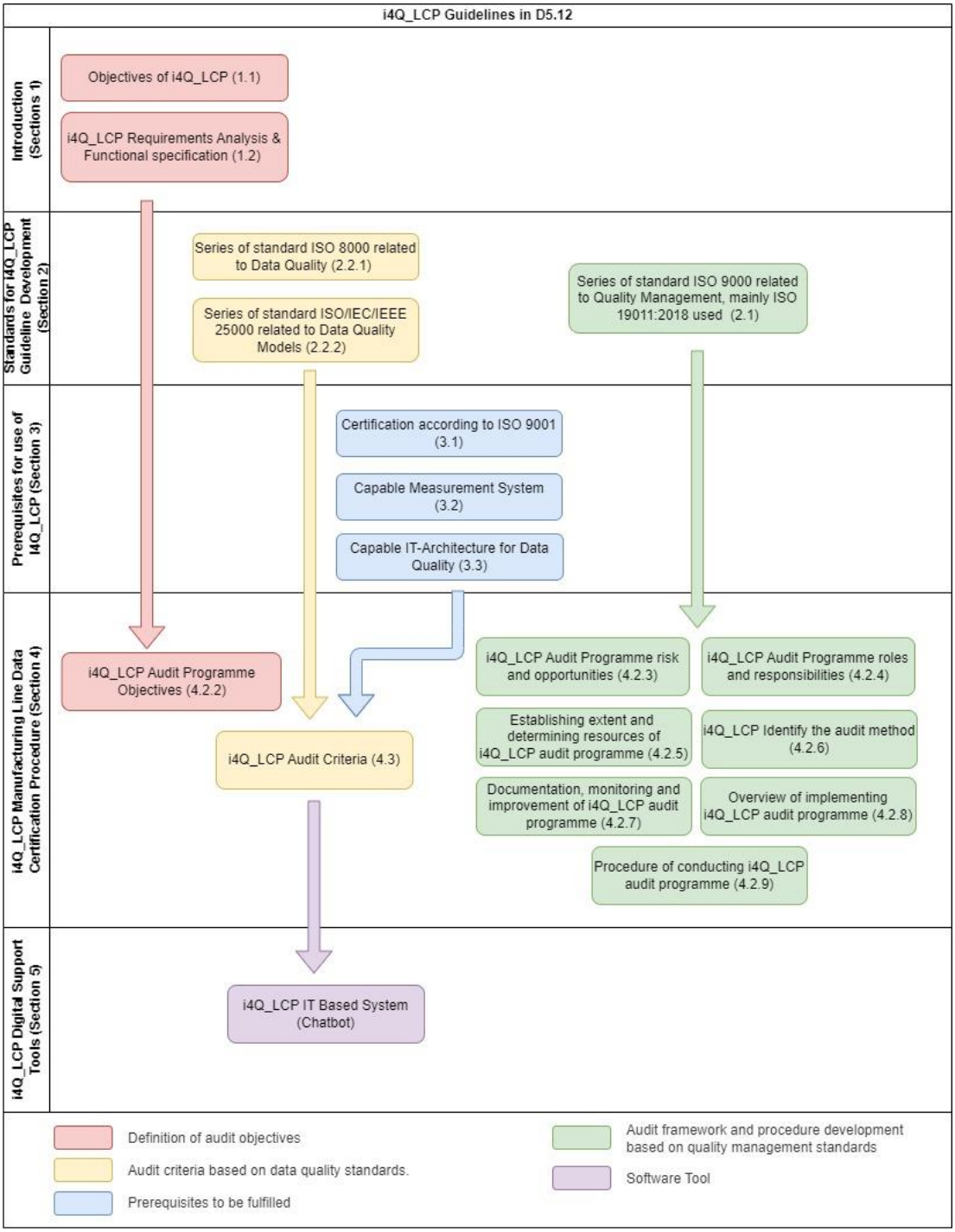

Figure 1. i4Q[LCP] Guidelines in D5.6 including references to sections in this document

Document structure

Section 1: The **Introduction** contains the objectives for the i4Q^{LCP} solution as well as its context in i4Q project including the status of fulfilment of stated requirements.

Section 2: In **Standards for i4Q^{LCP} Guideline Development**, the central guidelines for quality management and auditing management systems (ISO 19011:2018) are presented. Furthermore, relevant data quality standards including data quality models are introduced.

Section 3: The **Prerequisites for the use of i4Q^{LCP}** are presented including the certification according to quality management standards and the requirements for a capable measurement system in industrial manufacturing according to several standards. Also, a description of a capable IT-architecture is given as a prerequisite to use the data quality audit procedure successfully.

Section 4: In this section, the solution **i4Q^{LCP} Manufacturing Line Data Certification Procedure** is presented including the design and implementation of the audit framework and the audit criteria.

Section 5: The **i4Q^{LCP} Digital Support Tool** is presented in this section. It outlines the features of the Audit Advisor chatbot guiding users through the extensive questions to prepare for an audit.

Section 6: The **Conclusions** describe the outcome of this deliverable and the objectives of i4Q^{LCP} that have been achieved.

Appendix: The Appendix includes information and data on 4 subjects: Requirements and Functional Specifications of i4Q^{LCP}, Terms and definitions related to audits and data quality, and Audit Criteria Questions.

1. Introduction

The introduction contains the objectives for the presented solution and a summary of the preliminary work performed in WP1 and WP2 including the identified requirements and functional specification. In addition, the context to the i4Q project is outlined in relation to other i4Q WPs.

1.1 Objectives of i4Q^{LCP}

The objective of i4Q Manufacturing Line Data Certification Procedure (i4Q^{LCP}) is to develop a guideline which ensures data quality in smart manufacturing by providing recommendations for certification and audit procedures. The guideline is to be supplemented by a digitized IT- assistant to ensure the best usability for all auditors and users. The audit procedure for manufacturing resources (e.g., machine, cell, or manufacturing line) will be defined to assure that data generated by manufacturing processes is accurate and reliable. This procedure can be followed by internal or external auditors. The procedure describes the activities to be performed during audits as well as the elements of the manufacturing resources to be audited, the calibration equipment to be used, and the tests to be performed. It can also be used to supplement current quality certifications (i.e., DIN EN ISO 9000:2015–11 [20]) by adding the need of generating and ensuring data quality during manufacturing processes. The guideline consists of terms and definitions related to audit and data quality along with the structure and applicability boundaries within the manufacturing process. It also addresses the plan-do-check-act approach for planning, implementation, controlling, improvement, and documentation of data-driven qualification, reconfiguration, and quality control. Recommendations for process reconfiguration, audit strategies, certification, and regulation will be provided within the guide and related to the outcome of i4Q tasks T5.1 to T5.4.

This second deliverable D5.12 presents the final version of i4Q^{LCP} guideline development including a framework for the audit programme according to ISO 19011:2018 [6] and all audit criteria that have been developed based on ISO/IEC 25012:2008 [8] and ISO/IEC 25024:2015 [9]. The i4Q^{LCP} audit programme includes the logical sequence of activities to be performed during the audit procedure the audit criteria questions that guide the auditor through the audits of the DLC stages. The i4Q^{LCP} digital IT-assistant for audits and certification provides the audit criteria questions to users and enables them to prepare for an audit. It consists of a questionnaire part realized with MS Forms and a stand-alone chatbot that uses the guideline's content and workflows in its dialogs.

1.2 Context in i4Q Project

Data Quality is an essential aspect of data collection, transformation, and analytics within the industrial manufacturing environment; therefore, this guideline is a basic solution in the context of i4Q project. i4Q^{LCP} is a stand-alone guideline that can be used as a supporting tool for any i4Q pilot to ensure high data quality. Since the requirements collected in WP1 are imprecisely mapped and do not interface with other i4Q solutions, with the exception of i4Q^{QE}, and there are no defined requirements from developers other than TUB, there is a freedom for the development of this i4Q solution. As this solution is mainly understood as a guideline, no functional or technical requirements have been mapped. The chatbot of i4Q^{LCP} will use the same infrastructure as the assistant for i4Q QualiExplore (i4Q^{QE}) in Task 3.1.

Requirements and expectations defined by all involved partners as well as the functional specifications of the i4Q Solutions have been presented in i4Q Deliverable **D1.9 – Requirements Analysis and Functional Specification v2**[1]**.** The detailed description of the i4Q^{LCP} functional specification through Function Structure Diagrams (FSD) including the exact mapping of the requirements to the solution functions can be found in Appendix 1.

Table 1 shows the current progress in the implementation of the requirements.

ID	Type of requirement		Require-ment Title	Requirement definition (related to the i4Q solution)	Require-ment Coming from Solution	Compli-ance with require-ments
22TUB r1	Guide-lines req		Develop Audit Procedure	An audit procedure for manufacturing resources (machine, cell or manufacturing line) shall be developed to ensure the accuracy and reliability of data from the manufacturing process.	22-i4Q_LCP	100%
22TUB r2	Guide-lines req	Usability and Quality req	Knowledge Management	Knowledge shall be transferred to the employees through training according to guidelines and applied standards.	22-i4Q_LCP	100%
22TUB r3	Guide-lines req	Usability and Quality req	IT-Assistant	An IT-assistant should be developed to support audit teams to conduct the audit procedure	22-i4Q_LCP	100%

Table 1. Progress in the implementation of the i4Q^{LCP} requirements

Since i4Q^{LCP} is a guideline with a stand-alone software tool, there is no mapping to the i4Q Framework's architecture developed in WP2. Only the descriptions of stakeholders (D2.3) and tasks with corresponding roles (D2.4) could be referred to. According to **D2.3 – Report on Business Viewpoint**, i4Q^{LCP} affects all primary stakeholders, i.e., the internal team (D2.3, p. 27, Figure 8). The Production Team, Quality Team as well as the Maintenance Team and the Engineering Team could benefit from this solution and support its implementation. Specifically, there is an affiliation of i4Q^{LCP} to the Process Support Engineer, Processing Operator, Production Scheduler, Quality Manager, Quality Inspector, Maintenance Staff, Maintenance Service Planner, Process Support Engineer, and Data & Analytics Engineer. Secondary stakeholders may also be affected by this solution, for instance Standardization and Certification bodies if a standardization process occurs.

[1] Nowak-Meitinger, A. M.; Mayer, J. P.; Jochem, R. (2021). i4Q – *Deliverable 1.9 - Requirements Analysis and Functional Specification v2*. (Submitted 31-09-2021)

According to **D2.4** –Report on Usage Viewpoint, in which the guidelines have not been considered, i4Q^{LCP} could lead to a new task that involves auditing and ensuring data quality. Task 3.1 "Manufacturing Data Quality Strategy" contains the conceptual basis and an activity framework to manage data quality in production. It focuses on data quality factors and their management to maintain data quality. D3.1 covers the descriptions (guideline), while D3.2 provides a solution to store and organize data quality factor descriptions. i4Q^{LCP} will be evaluated and validated in i4Q WP6 by applying the audit procedure against specified and real-world scenarios and requirements at individual pilots.[2]

To ensure zero defects in the manufacturing line, T5.1 to T5.4 offer technical solutions that are designed to create new, more effective approaches and methodologies for process qualification as well as process reconfiguration and optimization using existing manufacturing data. i4Q^{LCP} takes into consideration the outputs of these solutions and give suggestions for further actions or tasks to be fulfilled in order to enhance the outcomes. For instance, i4Q^{PQ} and i4Q^{QD} mainly evaluate the process quality of manufacturing line with collecting and measuring the machine parameters and manufacturing conditions or data, i4Q^{LCP} guides about the standards for the measurement of these parameters and the measurement system [5], which need to be in placed to achieve good and reliable data quality which result in better accuracy of outputs of the solutions. The audit criteria questions developed in i4Q^{LCP} include the recommendation tasks that are derived from standards. The fulfilment of the criteria guarantees the certification of manufacturing process. This can be evident from the **Table 2** which show exemplary questionnaires from audit criteria of the data collection phase, that the quality of entered data can be verified and ensured. For instance, in T5.1 i4Q^{LCP} as a software solution, the end user will insert the upper and lower critical limits in order to calculate Cpk value which later can be verified with solution i4Q^{LCP} . Furthermore, prerequisites for usage of the guideline are described in detail in Section 3, based on the fulfilment of these prerequisites the user shall alter the audit strategies and decide whether i4Q^{LCP} to be used.

[2] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

Data Quality Characteristics	Description	Questions for Guideline	Answer	Task
Completeness	Completeness of data items of a record within a data file	Does your company use processes to ensure whether data of a relevant property is complete?	Yes	
			No	Processes to ensure the completeness of relevant properties shall be defined, implemented and documented.
Accuracy	Ratio of measurement coverage for accurate data	Does your company use processes to determine the accuracy of external data?	Yes	
			No	Processes to determine the data's accuracy shall be defined, implemented and documented.

Table 2. Exemplary audit criteria questions for T5.1

WP7 will disseminate and provide exposure for the i4Q^{LCP} and its outcomes in order to stimulate their adoption both inside and outside of the industrial sector. During i4Q internal standardisation workshops, i4Q^{LCP} has been identified as a high potential solution for standardisation activities. E.g., a CEN Workshop Agreement (CWA) could be developed in WP7 activities based on i4Q^{LCP} . In WP8, exploitation potential is identified. i4Q^{LCP} will be provided as an open-source solution.

2. Standards for i4Q^{LCP} Guideline Development

This section covers all relevant standards used to develop the i4Q^{LCP} guideline and audit procedure. The guideline and audit procedure developed within the i4Q project is based on standards from four main topic aeras,

- Quality Management Systems and related standards,
- Guideline for Auditing Management Systems,
- Data Quality, and
- Software Engineering.

ISO 9000:2015 Quality management systems – Fundamentals and vocabulary [20] introduces fundamentals and terminology used within all related standards. Further it provides the main principles of quality management to the user.

The *ISO 9001:2015 Quality management systems — Requirements and associated standards* [21] provides a general approach and requirements to establish a quality management system in companies while ISO 19011:2018 [6] covers all activities for auditing management systems.

The ISO 80xx ([11] – [19]) standard series addresses all data quality related activities and ISO 250xx [8] [9] standard series introduces quality to software development.

The i4Q^{LCP} guideline and audit procedure is based on relevant parts of the introduced standards. The standards are introduced in more detail in the following sections.

2.1 ISO 9000 Series and ISO 19011:2018

The most important series of standards in quality management is the cross-industry and internationally recognized DIN EN ISO 9000ff [23]. This standard family consists of the following standards that build a basis for the development of i4Q^{LCP}.[3]

The *ISO 9000:2015 Quality management systems – Fundamentals and vocabulary* is the first standard of the ISO 9000 series. This standard introduces the core concepts of quality management to the user. These fundamental concepts are defined as follows [20]:

- Quality
- Quality management system
- Context of an organization
- Interested parties
- Support
- People
- Competence
- Awareness
- Communication

[3] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

Additionally, the DIN EN ISO 9000:2015 –11 [20] introduces the quality management principles which are used throughout the ISO 9000 series:

- Customer focus
- Leadership
- Engagement of people
- Process approach
- Improvement
- Evidence-based decision making
- Relationship management

For each principle, the standard provides a clear statement, rationale, key benefits, and possible actions. Further the DIN EN ISO 9000:2015 –11 [20] introduces a general approach how to develop a quality management system (QMS) based on the mentioned principles and provides terms and definitions for all related standards. These vocabularies are used in i4Q[LCP] (see Appendix 2, **Table 12**).

As the fundamental standard for quality management systems with more than 1.2 million certified sites worldwide [22][4], the *ISO 9001:2015 Quality management system – Requirements* [21] provides all requirements for implementation and maintenance of quality management systems in companies on an organisational level. ISO 9001:2015 [21] defines the need to ensure measurement systems by setting requirements for them, such as documentation and standardization of measurement processes. This can be realized in combination with referenced standards like e.g., *ISO 15939:2017 Systems and software engineering – Measurement process* [5] and *ISO 10012:2003 Measurement management systems — Requirements for measurement processes and measuring equipment* [2], data quality is generated and maintained on its source, which is in context of the manufacturing industry the shopfloor with machines, manufacturing equipment, and implemented sensors. With a certification according to ISO 9001:2015 [21], a company ensures that all data quality related prerequisites are activities and processes are implemented, maintained, and continuously improved (see Section 3). Furthermore, ISO 9001:2015 [21] requires that internal audits shall be planned and conducted to ensure compliance with quality management requirements over time.

With the standard *ISO 19011:2018 - Guidelines for auditing management systems* [6], a complete document containing a core guideline for audit procedures and related information is provided to the industry. This standard outlines all principles for auditing, provides Best-Practice approaches for managing and conducting audits, and includes a general audit procedure according to the well-established Plan-Do-Check-Act circle (PDCA) and also competences and evaluation of the auditors. The developed i4Q[LCP] data quality audit guideline includes and uses the core elements of this standard.

[4] [22] https://www.iso.org/the-iso-survey.html

2.1.1 Principles of auditing

In order to be effective and provide insightful information that an organisation can act upon, ISO 19011:2018 [6] defines seven principles of auditing that should be fulfilled:

1. *Integrity* is seen as the foundation of professionalism, so the auditor should execute their tasks ethically, responsibly, honestly, and unbiased. They should also be aware of their own competencies and limitations and be sensitive to any influences on their judgement.
2. *Fair representation*, as auditing principle is the obligation to report the audit evidence, findings, and conclusions truthfully, accurately, objectively, timely, completely, and clear.
3. Furthermore, auditors should perform their work with *due professional care*, which includes diligence and the ability to judge any audit situation in a reasonable manner.
4. *Confidentiality* concerns the security of information: Auditors should properly handle sensitive or confidential information they acquire during the performance of their work.
5. For an audit to be impartial and objective auditors also should be *independent* of the activity being audited, if possible.
6. The sixth auditing principle is the *evidence-based approach*. The audit evidence should be verifiable so that the audit conclusions are reproducible and reliable.
7. Lastly, audits and audit programmes need to consider risks and opportunities, to ensure that the focus lies on aspects that are significant to the client (*risk-based approach*).[5]

2.1.2 Managing and conducting an audit

The audit procedure structure is defined in the form of the Plan-Do-Check-Act (PDCA) method, which includes all the necessary activities starting from planning, implementing, managing, monitoring, and controlling the audit for data quality. The main aspects of this work can be found in **Figure 2**. Important steps are the preparation of the audit programme, including the definition of audit objectives, audit scope, audit schedule and audit criteria. Also, risks should be considered in the "Plan" phase and the documentation procedure defined, e.g., physical, or digital checklists, audit sampling details and audio-visual information. In the "Do" phase, the audit is conducted and evidence for the fulfilment of audit criteria provided. In the "Check" phase, monitoring of the audit programme takes place, and in the "Act" phase, an evaluation with suggestions or recommendations should be made. This approach is used to create the framework of the i4Q[LCP]. Audits can be managed and performed as first-party, second-party or third-party audits. The first-party audit is a purely internal audit. The second-party audit is performed by an external interested party, e.g., a customer auditing a supplier. The third-party audit is a certification and/or accreditation audit where a certificate is given for specific standards and regulations [6][23].

[5] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

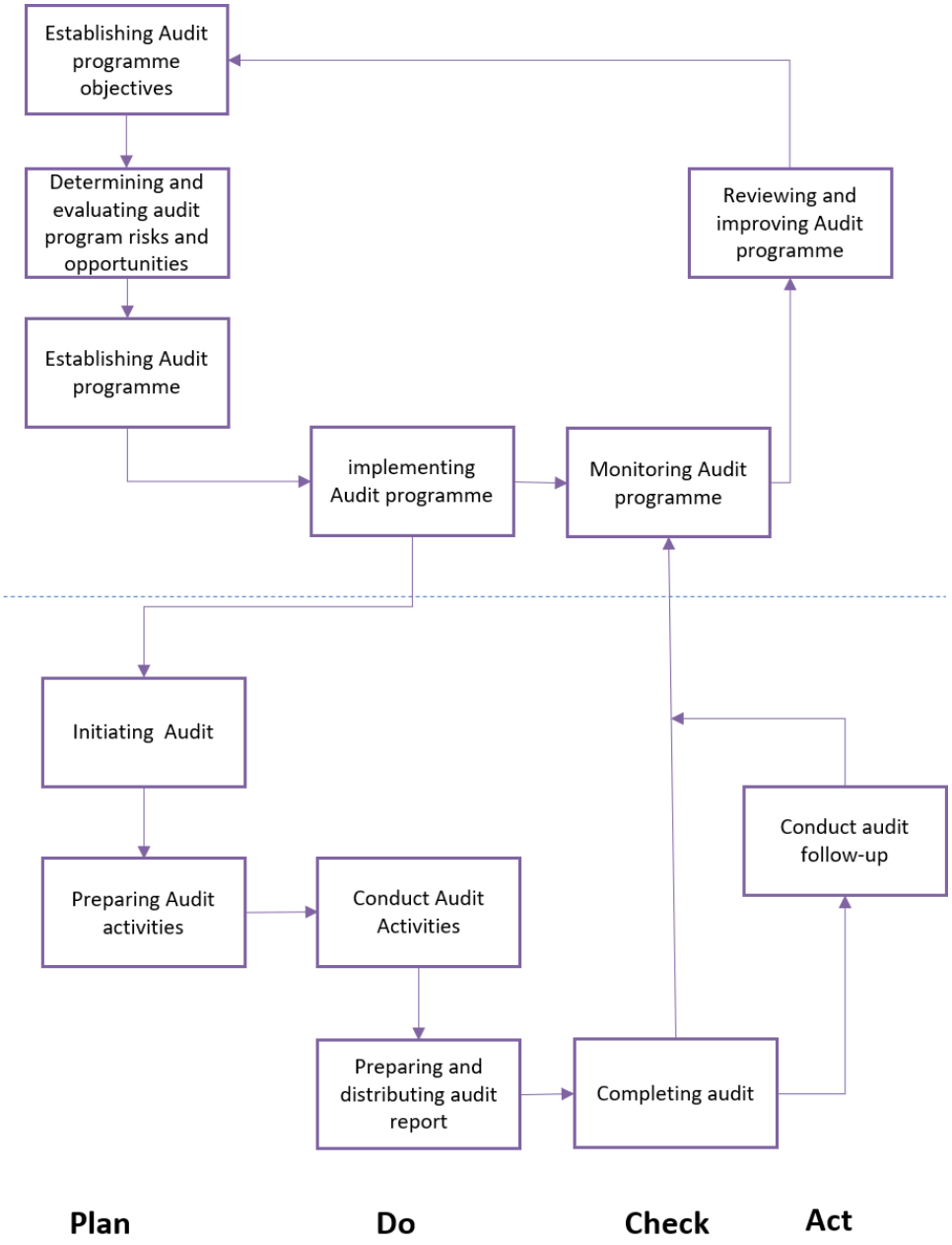

Figure 2. Managing and audit programme (upper part) and conducting an audit (lower part) according to PDCA method [6]

One of the most important activities in the audit process is the collection and review of information to provide evidence for the assessment of audit criteria. This conformity assessment is required to derive audit conclusions. Audit methods should be selected according to audit requirements. **Figure 3** shows such a typical process, which starts with the information source providing information through appropriate sampling to obtain audit evidence and then it is

evaluated against the audit criteria. In this way, audit findings are obtained, and the review is performed, thus finally leading to the audit conclusions [6][6].

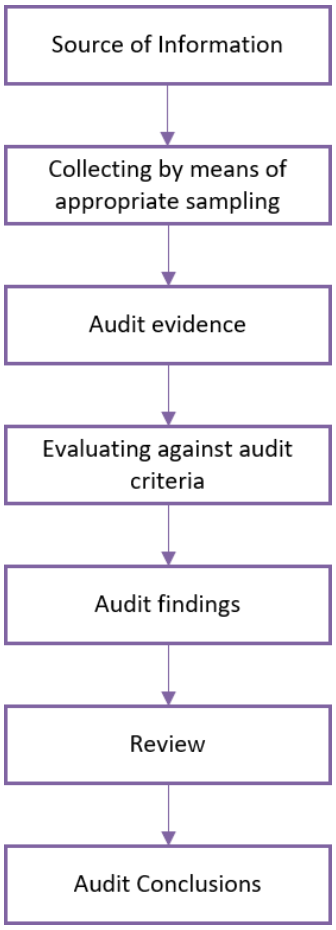

Figure 3. Typical process of collecting and verifying information [6]

2.1.3 Auditors

ISO 19011:2018 [6] describes different aspects of auditor competence and distinguishes between personal behaviour and knowledge and skills. In terms of *personal behaviour*, auditors should exhibit professionalism which includes ethical, truthful, open-minded, diplomatic, perceptive, versatile, tenacious, decisive, culturally sensitive, and collaborative behaviour. *Knowledge and skills* include generic knowledge and skills of audit principles (see section 2.1.3), processes and methods, management system standards, the auditee organization, and its context as well as applicable statutory and regulatory requirements. Additionally, discipline and sector-specific competence of auditors are necessary. This includes fundamentals of the auditee's discipline and sector, application of discipline and sector-specific methods, techniques, processes and practices and knowledge of management system requirements and principles. To evaluate auditor competence, both qualitative and quantitative criteria are used. Auditor evaluation methods can

[6] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

be reviews of records, feedback, interviews, observations, tests, and post-audit reviews. Auditors of management systems should be able to plan and organize the work effectively in order to perform the audit within the agreed time schedule. They should collect information through effective interviewing, listening, observing, and reviewing documented information, including records and data. Furthermore, auditors should understand the appropriateness and consequences of using sampling techniques for auditing. Also, they should verify the relevance and accuracy of collected information and assess and/or confirm the sufficiency and appropriateness of audit evidence to support audit findings and conclusions. Also, the documentation and reports of the audit activities and audit findings are part of the work of an auditor [6][7] The exact competencies required for i4Q[LCP] are described in Section 4.

2.2 Data Quality relevant Standards

With the increasing digitization of the manufacturing industry as well with use of modern technologies for data analytics and data driven decision making, data quality is considered more crucial than ever. Especially for the solutions developed with in the i4Q project high data quality is needed to ensure full functionality and use to full potential.

2.2.1 ISO 8000 Data Quality series

An important set of standards related to data quality is the series of standard ISO 8000. This series of standards defines characteristics of data quality, provides frameworks for improving data quality and covers industrial data quality characteristics throughout the product life cycle. Therefore, the ISO 8000 series is also included as source for information for the development of i4Q[LCP] guideline and audit procedure. The ISO 8000 series provides the basic principles of information and data quality and general approaches to ensure data quality from different perspectives and within several main topic fields. The standards mainly used for the development of i4Q[LCP] are described below.[8]

ISO 8000-2:2020 Data quality – Part2: Vocabulary provides clear definitions of all relevant terms in context with data quality. A table of used terms in i4Q[LCP] is provided in Appendix 3.

ISO 8000-8:2015 Data quality — Part 8: Information and data quality: Concepts and measuring introduces measuring concepts for data quality. Without these concepts, data quality cannot be managed and therefore, not be verified, validated and improved. This standard considers data from various sources. Defined data quality requirements to ensure syntactic, semantic and pragmatic data quality are used within i4Q[LCP].

The standard *ISO 8000-61:2016 Data quality — Part 61: Data quality management: Process reference model* specifies a process for data quality management with a defined purpose, outcomes and activities. This reference model needs to be conducted in order to assure data

[7] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

[8] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

quality within a company. The development of the audit procedure of i4Q[LCP] is based on this process. Detailed structure of the data quality management processes in ISO 8000-61:2016(E) is shown in **Figure 4.**

With the standard *ISO 8000-63:2019 Data quality — Part 63: Data quality management: Process measurement* [15], a three-level procedure to measure processes is introduced. The three levels consist of process, inspection, and a maturity model and are integrated into process model for assessment of current data quality state within a company (**Figure 5**). This procedure is used within i4Q[LCP] to assess the current level of data quality within the company and is part of the digital tool.

ISO/TS 8000-81:2021 Data quality — Part 81: Data quality assessment: Profiling [18] provides and specifies a procedure for data profiling for data quality assessment applicable to data sets structured in tables and columns.

ISO/TS 8000-150:2022 Data quality — Part 150: Master data: Quality management framework Roles and responsibilities [19] provides key considerations for organizations to assign appropriate roles and responsibilities for data quality management. This is considered as one key element to ensure data quality within companies. The standard defines implementation roles and documentation, a framework and a functional model for roles and responsibilities, and exemplary deployment scenario based on the earlier mentioned ISO 8000-61:2016 [13]. The introduced roles are also used within i4Q[LCP].

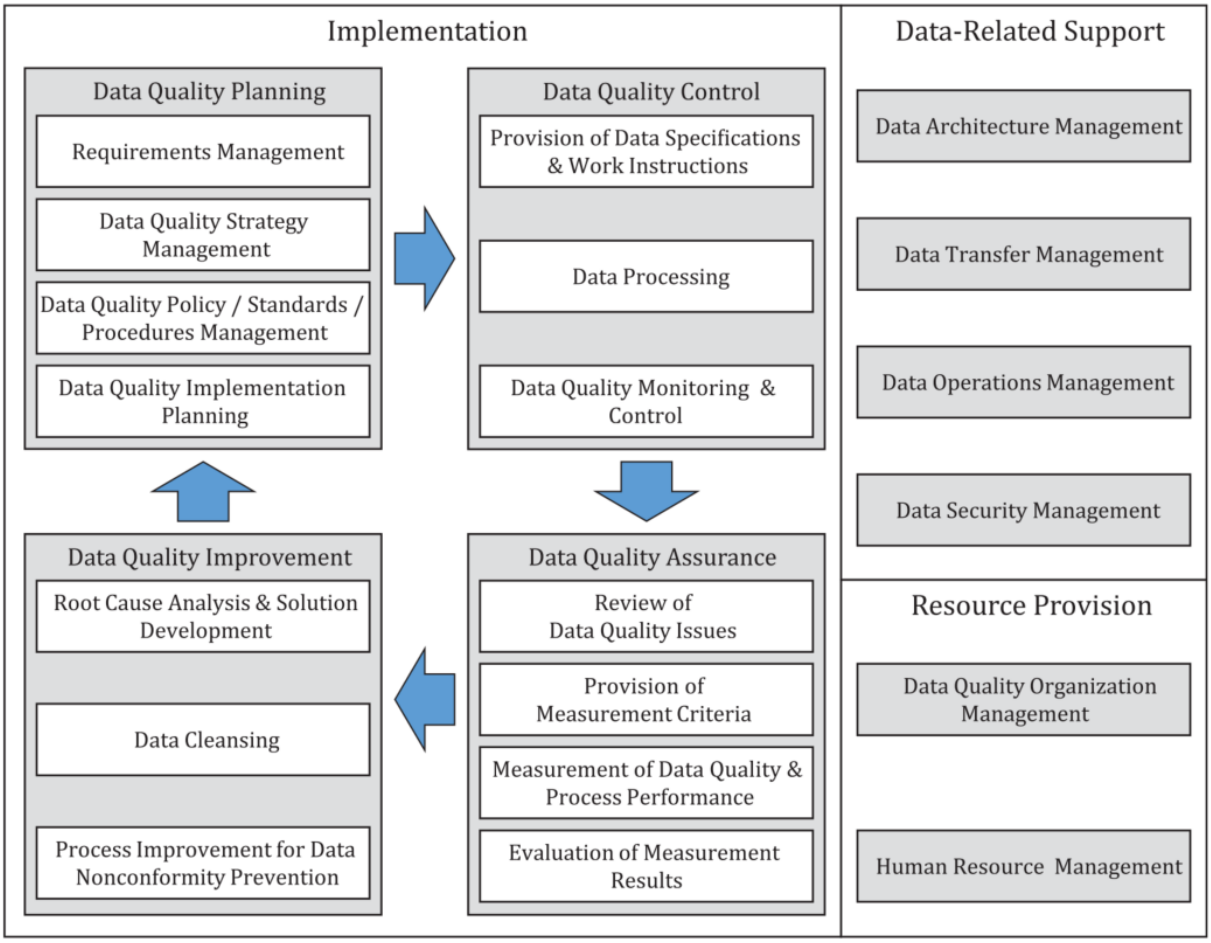

Figure 4. Detailed structure of the data quality management processes in ISO 8000-61:2016 [13]

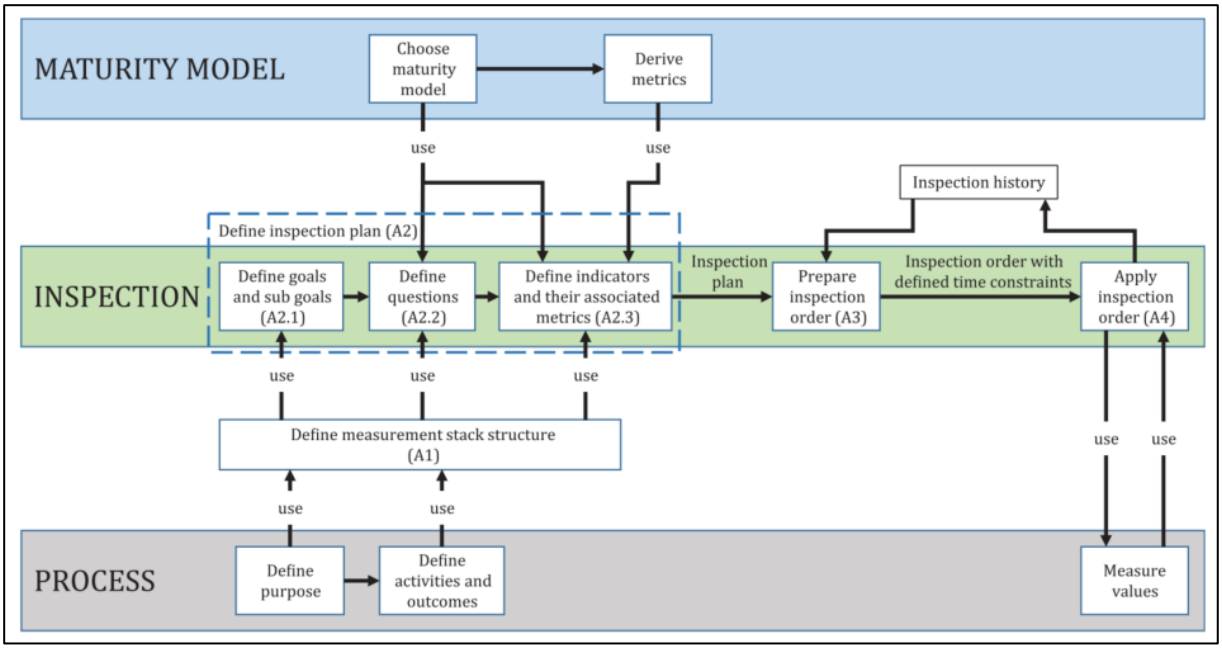

Figure 5. Relationship between inspection plan, maturity model and process according to ISO 8000-63:2019 [15]

2.2.2 ISO IEC 25000 System and software engineering Series

Coming from the area of system and software engineering the ISO IEC 25000 series presents a complete set of standards to ensure quality through use of defined Requirements and evaluation processes. The provided data quality requirements and models are incorporated into the i4Q[LCP] guideline and audit procedure. This applies also to the *Data life cycle model* (DLC) containing all relevant phases for processing and handling data, the assigned requirements, and their measurements.

The *ISO/IEC 25012:2008 Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model* [8] introduces a *data quality model* consisting of two, inherent and system depended points of view. The first describes the inherent data quality and its data quality characteristics. Therefore, it addresses data quality within data. The second focuses on system-dependent data quality. Because data is processed and handled within systems, computer systems and software, it is affected by both. The standard provides a detailed overview of *data quality model* characteristics (**Table 3**).

Characteristics	Data quality	
	Inherent	System dependent
Accuracy	X	
Completeness	X	
Consistency	X	
Credibility	X	

Currentness	X	
Accessibility	X	X
Compliance	X	X
Confidentiality	X	X
Efficiency	X	X
Precision	X	X
Traceability	X	X
Understandability	X	X
Availability		X
Portability		X
Recoverability		X

Table 3. Data Quality Model Characteristics [8]

The generic DLC introduced in the *ISO/IEC 25024:2015 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of data quality* [9] is referred as a basic framework for the i4Q^{LCP} development. It covers all essential Stages of the *Data Life Cycle* like *data design, data collection, external data acquisition, data integration, data processing, presentation, other use, storage* and *delete.* These stages cover all data processing and handling activities defined in the i4Q pilot use cases. Because of their generic design, these stages are variable as per pilot's need and nature of internal system. The DLC stages covered in i4Q^{LCP} as well as the target entities and properties are described in Section 4.3.

3. Prerequisites for the use of i4Q^{LCP}

As data from manufacturing lines is the source of information for almost all analytical i4Q solutions and no procedure exists to audit the Data Life Cycle, i4Q^{LCP} is developed to fill this gap. For the use of i4Q^{LCP} guideline there are prerequisites and activities a company must conduct. [9] This section highlights the prerequisites that should be fulfilled to ensure the use of i4Q^{LCP}. The main aspects to be fulfilled are the implementation of a quality management system according to ISO 9001:2015 [21], a capable measurement system and a capable IT architecture for data quality. These three aspects are described in the following subsections and the corresponding requirements of these standards are part of the pre-audit criteria questionnaires.

3.1 Certification according to ISO 9001

In the manufacturing industry the use of *ISO 9001:2015 Quality management systems — Requirements and associated standards* [21] is seen as best practice. Therefore, as a basis, prerequisites and requirements coming from this widely used and commonly applied standard need to be fulfilled. The organization shall have a working quality management system and be certified according to ISO 9001:2015 [21]. This requires established basic elements of measurement systems to be used in the organization. These measurement systems are documented and subject to monitoring, which are checked at specified time intervals.

While ISO 9001:2015 [21] only briefly mentions the use of a suitable measurement system and the need for documentation and verification of the results, standards like ISO 10012:2003 [2] and ISO 15939:2017 [5] should be used for more detailed requirements that have to be fulfilled to achieve a capable measurement system.

3.2 Capable Measurement System

Since the main source of manufacturing data within the i4Q project is the manufacturing process, high quality machine data collected by sensors have a significant impact on most of the analytical i4Q solutions. In reference to ISO 9001:2015 [21], ISO 10012:2003 *Measurement management systems — Requirements for measurement processes and measuring* [2], and ISO/IEC/IEEE 15939:2017 *Systems and software engineering — Measurement process* [5], a suitable measurement system shall be implemented to ensure high quality data directly from the manufacturing process.

The ISO 10012:2003 [2] defines a clear approach for introduction of a measurement system, defined as metrological function, to a company in context to fulfil customer expectations and, therefore, quality. This approach covers the following aspects[10]:

[9] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

[10] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

- General requirements
  - Ensure fulfilment of use case specific metrological requirements for the used measurement system.
- Management responsibility
  - Focus on definition of metrological function, customer requirements, quality objectives, and management review to ensure continual adequacy, effectiveness, and suitability of the measurement system.
- Resource management
  - Allocation of human resources in terms of responsibilities and training
  - Provision of information resources like e.g., standardized procedures, needed software, providing required records, and clear identification of measurement system
  - Provision of material resources in terms of the measurements system itself and documentation of environmental conditions
  - Outside suppliers have to provide a clear and specific documentation of products and services for the measurement system
- Metrological confirmation and realization of measurement processes
  - Confirmation of fulfilment of all metrological requirements for the used measurement system and its characteristics e.g., range, bias, repeatability, etc. by calibration, recalibration, and/or equipment verification.
  - Definition of intervals for reviewing the metrological confirmation
  - Introduction of measures to protect the measurement system against misuse and tampering
  - Establishment of a documentation process for the metrological confirmation by authorized persons
- Measurement process
  - Design and definition of measurement process according to requirements from customers, organization, and statutory and regulatory requirements
  - Realization of the measurement process under controlled conditions
  - Confirmation of compliance to requirements for the measurement process through records and documentation
- Measurement uncertainty and traceability
  - Estimation and recording of all possible uncertainties
  - Inurement of traceability of all measurement results to SI unit standards
- Measurement management system analysis and improvement
  - Planning of audit and monitoring measures to insure suitability and effectiveness of the used measurement system
  - Control measures to deal with any kind of nonconformities
  - Planning of future continual improvement, corrective, and preventive measures

Every phase of the described approach requires documentation to provide evidence for performing audit activities according to ISO 10012:2003 [2] and ISO 9001:2015 [21]. **Figure 6** shows the Metrological Confirmation Process according to ISO 10012:2003 [2].

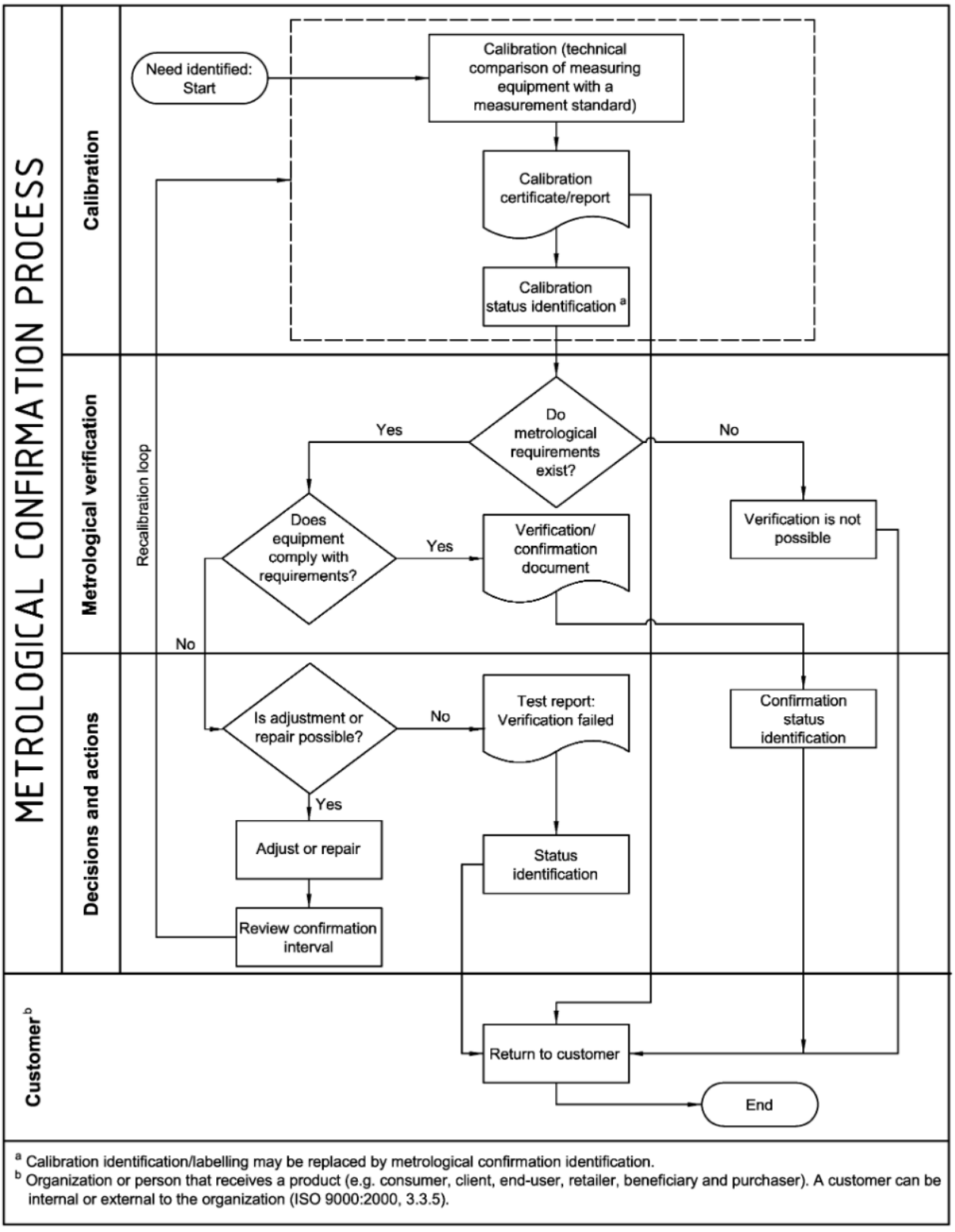

Figure 6. Metrological Confirmation Process according to ISO 10012:2003 [2]

The ISO/IEC/IEEE 15939:2017 *Systems and software engineering — Measurement process* introduces a top-level procedure to establish measurement processes in the industrial environment which includes a measurement system (**Figure 7**). The requirements for a suitable

measurement system are described in ISO 10012:2003 [2]. The standard covers the following aspects[11]:

- Purpose of the measurement process is *Data collection, Data analysis and reporting objective data* for management and demonstration of compliance to defined quality requirements.
- Outcomes as defined in ISO/IEC/IEEE 15288:2015 *Systems and software engineering — System life cycle processes, 6.3.7.1* [4] and ISO/IEC/IEEE 12207:2017 *Systems and software engineering - Software life cycle processes* [3]
- Main tasks:
  - Establish and sustain measurement commitment as part of project planning and assessment and control products:
    - Definition of requirements for measurement
    - Assignment of resources
  - Prepare for measurement:
    - Development of the measurement strategy
    - Description of organizational characteristics which are relevant to measurement
    - Provision of identified and prioritized information needs based on the organization
    - Definition of measurement which fulfils the needs,
    - Definition of data collection, analysis access and reporting process
    - Definition of evaluation criteria for information item and measurement process
    - Identification and planning of enabling systems and/or services
  - Perform measurement
    - Integration of all needed procedures into relevant processes
    - Collection, storage, and verification of data
    - Analysis of data and development of information items
    - Recording of results and report to measurements to user
  - Evaluation of measurement by quality assurance in compliance to ISO/IEC/IEEE 15288:2015 [4]
    - Evaluation of information products and the measurement process
    - Identification of potential improvements

  Each main task consists of several subtasks and activities which must be performed in order to establish a standardized and stable measurement process.

[11] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

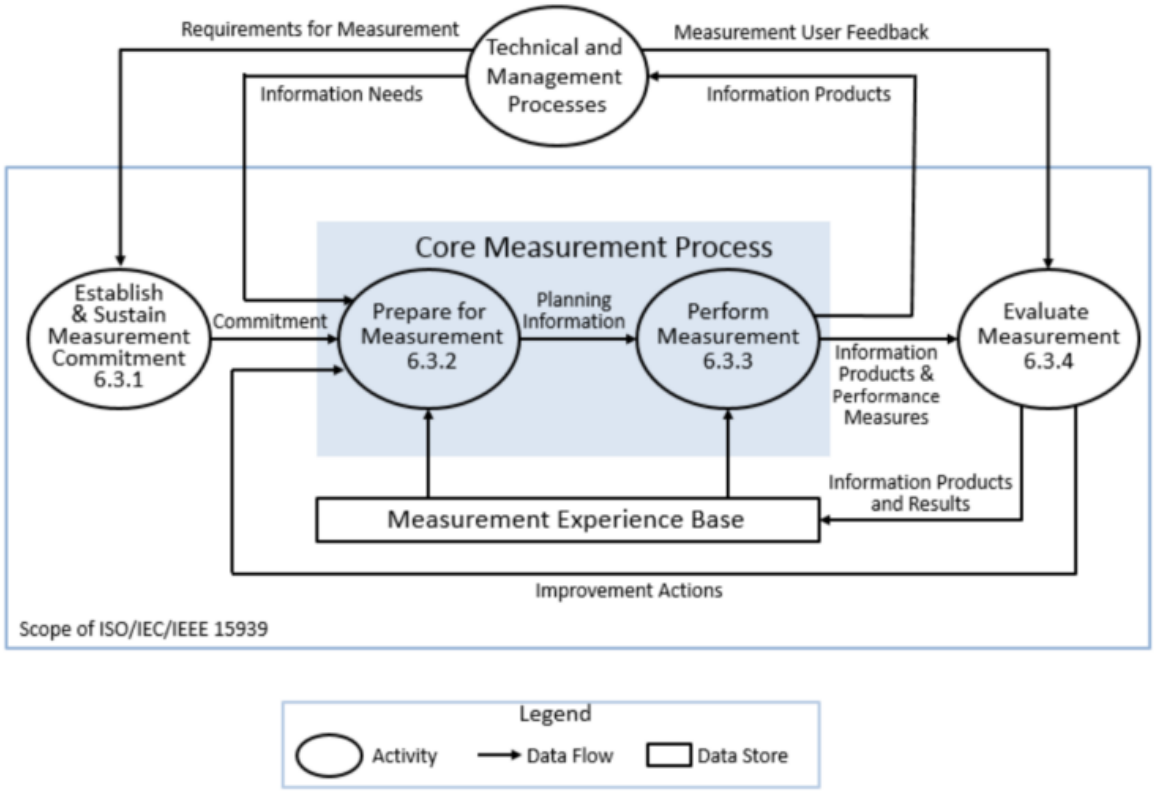

Figure 7. Measurement process model (ISO /IEC/IEEE 15939-2017(E))

As the described standards are commonly used in the manufacturing industry and most companies are certified according to ISO 9001:2015 [21], it is assumed that all activities to achieve stable and valid measurements are performed before starting the audit procedure i4Q[LCP].

For "off-the-shelf" measurement systems, services, and software it is assumed they are implemented and used as intended by the provider of the product. Therefore, related aspects and steps of the i4Q[LCP] solution can be skipped and the user can refer to certificates and/or regulations applicable to the used product provided by the manufacturer. This is compliant to ISO 10012:2003 [2][12].

3.3 Capable IT-Architecture for Data Quality

A capable, i.e., error-free and functional, IT architecture is a prerequisite for data quality. Processes within the architecture shall run without errors, and best practices shall be implemented and standardized. The standard ISO IEC 33020:2019 *Information technology - Process assessment - Process measurement framework for assessment of process capability* [10] is suitable for assessing IT processes.

In ISO/IEC 33020:2019 [10] the process attributes are organized into process capability levels, ranging from Incomplete (Level 0) to Innovating (Level 5). These process capacity levels are used to organize the process attributes. The outcome can also be described in terms of the ratings for

[12] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

each process' capability levels that were attained. An organization's capability level rating only indicates that it is capable of conducting its procedures at that level; it does not ensure that it will do so. Process capability levels are defined in ISO/IEC 33020:2019 [10] as follows:

- Process capability Level 0: *Incomplete process* – The process is not implemented or fails to meet its objectives. There is little or no evidence of systematic achievement of the process goal at this level.
- Process capability Level 1: *Performed process* – The process performance attribute measures how effectively the process purpose is fulfilled.
- Process capability Level 2: *Managed process* – The previously described *Performed process* is now managed (planned, monitored, and adjusted), and its documented information is established, controlled, and maintained appropriately.
- Process capability Level 3: *Established process* – The described *Managed process* is implemented using a defined process that is both assured and constantly improved.
- Process capability Level 4: *Predictable process* – The previously mentioned *Established process* is carried out predictably. Quantitative management requirements are identified, measurement data is collected and analysed, and assignable causes of variation are identified. To address assignable causes of variation, corrective action is taken.
- Process capability Level 5: *Innovating process* – The described *Predictable process* is continuously improved in order to respond to changes via identified innovative approaches for process innovation.

The process capability level is attained using scale levels and ratings. The ratings of the levels are divided into the two categories of "Largely" and "Fully". Process performance, performance management, documented information management, process definition, process deployment, process assurance, quantitative analysis, quantitative control, and process innovation are the attributes used to rate each scale level [10].

To use i4Q^{LCP} certification guideline, the process capability should be achieved to level 3 – *Established process* or higher to describe the process measurement requirements. The managed process (planned, monitored and adjusted) in level 2 shall be implemented using a defined process that is assured and continuously approved.

The attributed required to establish and maintain a standard process at level 3 include according to ISO/IEC 33020:2019 [10]:

1. Required inputs and expected outputs should be determined for the standard process.
2. Certain attributes like sequence withing the process, interaction, roles, competences, responsibilities, authorities and resources should be determined in a standard process.
3. Knowledge which is necessary for the operation of standard process is determined and maintained (knowledge source can be internal or external).

The attributes required for the deployment and assurances of the process include according to ISO/IEC 33020:2019 [10]:

1. Education, training and experience of the person(s) performing process should be competent.
2. Retaining documented information and its availability to ensure that the standard process is continuously improved based on identified needs and opportunities.

3. A strategy for establishment is managed, risks of the process are determined and evaluated.
4. Any non-conformity should be acted upon based on its nature effect and tracked to closure.

If level 3 is achieved and the attributes for establishing, deploying and assuring the process are met, a sufficiently capable IT architecture and thus the prerequisite for i4Q^{LCP} is fulfilled.

4. i4Q^{LCP} Manufacturing Line Data Certification Procedure

In this section, the framework of i4Q^{LCP} Manufacturing Line Data Certification Procedure is developed according to ISO 19011:2018 [6]. This is aligned to the theoretical basics explained in section 2.1. The complete guideline is presented in the following sections 4.2 and 4.3.

4.1 How to use / manual of i4Q^{LCP}

This section describes how to use the guideline in the simplified steps as shown in **Figure 8.** The user can start the process by selecting a particular cell or machine for scope of the audit and define the basic objectives of audit or audit program. Each audit program has set objectives, such as auditing individual DLC phases, therefore user must choose the desired DLC phase as mentioned in Sections 2.2.2 and 4.3.2, which is to be audited. The objectives shall be defined and checked as described in Sections 1.1 and 4.2.2. Once the objectives are defined and match the general objectives of i4Q^{LCP}, the next step of checking prerequisites can be done or in case objectives are not met, i4Q^{LCP} shall not to be used. If the prerequisites mentioned in Section 3 are not fulfilled by user, the necessary actions must be taken, and the prerequisites shall be fulfilled. In the next step, the audit framework shall be defined which includes formation of audit team and implementation and conducting an actual audit. The IT-based tool for i4Q^{LCP} is to be used as an assistance for the auditee to prepare and conduct the audit. The audit criteria mentioned in section 0 shall be assessed using the digital tool (Section 5). Once the assessment is completed with the digital tool, a detailed report will be generated and can be further used as described in 4.1.9. **Figure 8** shows that the steps where IT-tool assists user during audit preparations. It allows to choose the DLC stages, answer pre-questions related to prerequisites and evaluate the audit criteria. Colours for each step refer to those in **Figure 1**.

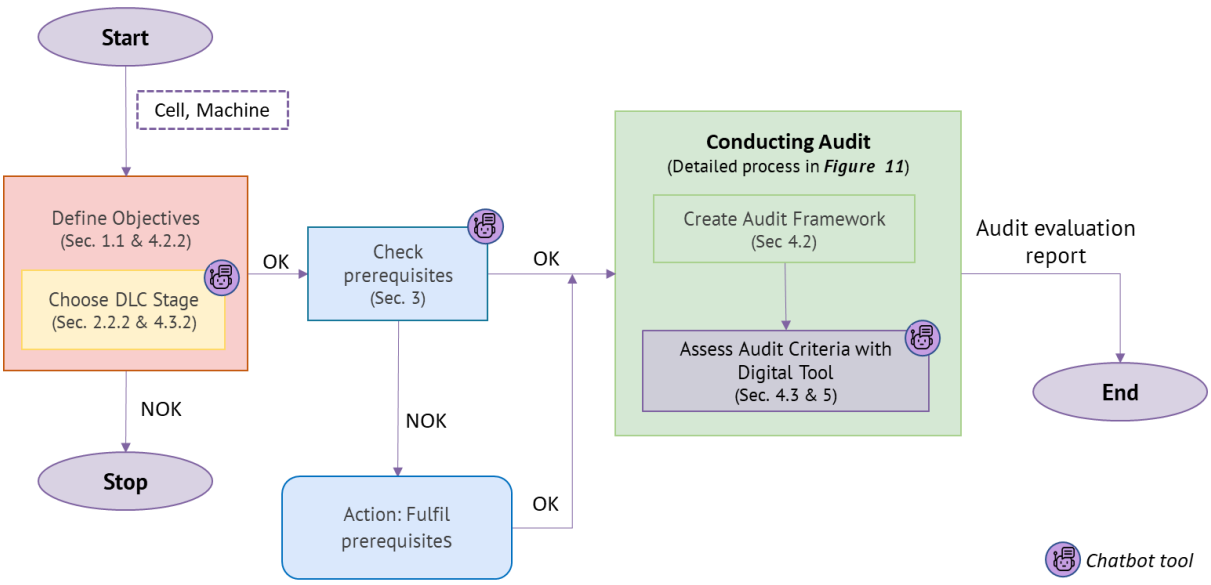

Figure 8. Steps for usage of i4Q^{LCP}

4.2 Procedure and framework of i4Q[LCP]

This section presents the procedure and framework of i4Q[LCP] to define the scope, objectives, extent, and other important aspects of the guideline. This framework is mainly based on ISO 19011:2018 [6] and has been published in the first version of this deliverable [32].

4.2.1 i4Q[LCP] General audit programme conditions and scope [13]

The audit programme developed in solution i4Q[LCP] includes the certification procedure related to auditing *data quality* only. As this is a new guideline, only the data quality in the DLC of the industrial environment is audited. No other management systems will be considered at the same time. Nevertheless, other standards are referred to in the respective parts of the guideline.

To ensure the usability of the audit procedure, it is designed for the use in manufacturing environment only. Auditees should focus on machines, cells, or manufacturing lines. The procedure will describe the logical sequence of the activities to be performed and data of the manufacturing resources will be audited, e.g., data coming from sensors, controls, or software. The procedure might be transferable to other use cases.

If parts or components of the data life cycle process are outsourced, auditees should ensure that all information and documents are available to assess the processes and systems. Experts for the respective systems may be required to attend the audit. To ensure confidentiality, documents should be uploaded in a safe mode and/or on local devices/servers.

Auditors and auditees shall act with integrity and not be influenced to ensure reliable audit results. Also, parts of the data life cycle or system elements that carry higher inherent risk and/or lower performance should be focused on to effectively improve the processes and overall data quality of the system. The roles and responsibilities of the individual(s) leading the audit programme are described in Section 4.2.4, to define the prerequisites for persons suitable for conducting the audit programme.

Other important information about the audit programme that is necessary for the effective and efficient conduct of the audit within the specified time frame is provided in the following sections. They describe, for instance, the objectives of the audit programme, the associated risks and opportunities and the measures taken to address them, the audit criteria, the audit methods, the criteria for selecting the members of the audit team, and the relevant documented information. Some of this information may vary depending on the circumstances and prerequisites of the auditee. In particular, the scope (extent, limits, and locations) and schedule (number, duration and frequency) of audits should be determined depending on the use case. For example, this could vary depending on locations and new machines, their reconfigurations, new batches, or recurring audits such as monthly or yearly. In addition, the implementation of the audit programme should be monitored and measured to ensure its objectives have been achieved.

[13] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

Finally, it is important to emphasize that i4Q^{LCP} provides a manufacturing line certification procedure for *internal audits*, also called *first party audits*. The certification process is not accredited (yet). Nevertheless, it gives a clear overview of the status of data quality in the entire process of data generation and handling in auditee's manufacturing environment. If all requirements for a high data quality measurement system are met and thus a sufficiently high data quality has been measured/assessed by the IT assistant, an internal certificate is provided.

There are three types of quality audits: system, process, and product audits [23]. In the case of i4Q^{LCP}, an audit procedure is developed to ensure the data quality of a specific manufacturing process.

4.2.2 i4Q^{LCP} Audit programme objectives [14]

Every audit programme has its specific objectives such as auditing single DLC phases related to a specific machine, cell, or manufacturing line. This must be defined by the use case with its auditors and auditees. The general audit programme objectives of i4Q^{LCP} are defined as follows:

- Ensure data quality of all systems and processes that are related to the DLC in the manufacturing environment.
- Measure maturity of data quality in the audited system.
- Identify opportunities for the improvement of data quality of the system and improve data quality where necessary.
- Evaluate the capability of the auditee to determine risks and opportunities related to data quality in the defined DLC and to identify and implement effective actions to address these risks and opportunities.
- Assess conformity to all relevant data quality requirements defined in the audit criteria.
- Determine the continuing suitability, adequacy, and effectiveness of the auditee's data quality system.
- Support ensuring a high data quality infrastructure to enable effective use of predictive methods such as provided in i4Q solutions of tasks T5.1 to T5.4, i.e., i4Q^{PQ}, i4Q^{QD}, i4Q^{PA} and i4Q^{LRT}.
- Provide an internal certificate to verify and compare maturity of data quality.

4.2.3 i4Q^{LCP} Audit programme risks and opportunities

Determining and evaluating the risks and opportunities of an audit programme is an important step in establishing a robust audit programme. To avoid the risks identified, good audit practice plans should be developed for each audit case. **Table 4** outlines general risks and associated measures. The list has been completed compared to D5.6. More specific risks can be applied to specific use cases of the auditee.

[14] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

Risks associated with the following	Risk	Measure
Planning		
Audit objectives	Objectives not met, not defined	Describe objectives clearly and easily understandable, adjust procedure if necessary
Extent of the audits	Misplanning	Test cases in i4Q project WP6, adjust procedure if necessary
Number of the audits	Misplanning (insufficient evidence of audit reports)	Test cases in i4Q project WP6, adjust procedure if necessary
Duration of the audits	Misplanning (more time required than planned)	Test cases in i4Q project WP6, adjust procedure if necessary
Locations of the audits	Misplanning (Change in location or workstation on the spot)	Test cases in i4Q project WP6, adjust procedure if necessary
Schedule of the audits	Misplanning (Unavailability of resources. For instance, audit team members, equipment)	Test cases in i4Q project WP6, adjust procedure if necessary
Resources		
Time for developing the audit programme or conducting an audit	Allowing insufficient time for developing the audit programme or conducting an audit	Allocating sufficient time frame to develop the audit programme or conducting audit to achieve desired audit objectives.
Equipment for developing the audit programme or conducting an audit	Allowing insufficient equipment for developing the audit programme or conducting an audit	Providing necessary equipment to develop the audit programme or conducting audit to achieve desired audit objectives.
Training for developing the audit programme or conducting an audit	Allowing insufficient training for developing the audit programme or conducting an audit	Preparing and providing training for developing the audit programme or conducting an audit.
Selection of Audit Team		
Competence to conduct audits effectively	Insufficient overall competence to conduct audits effectively	Matching the level of competence of the audit team to the level of competence needed to achieve the audit objectives
Communication		
Internal communication process	Ineffective internal communication process/channel	Internal communication matrix to be formed where in all the members are included and communicated.

External communication process	Ineffective external communication process/channel	External communication matrix to be formed where in all the partners are included and communicated.
Coordination & Security		
Coordination of the audits within the audit programme	Ineffective coordination of the audits within the audit programme	Effective coordination plan of audits to be developed and implemented within the audit programme.
Considering information security and confidentiality / trustworthiness	Not considering information security and confidentiality	Proper binding confidentiality agreement to be presented and signed by all stakeholders in order to ensure information security.
Control of documented information		
Necessary documented information required by auditors and relevant interested parties	Ineffective determination of the necessary documented information required by auditors and relevant interested parties	List of required documented information should be prepared at the beginning of the audit to ensure information required by auditors and relevant interested parties
Protect audit records to demonstrate audit programme effectiveness	Failure to adequately protect audit records to demonstrate audit programme effectiveness	Document or information control plan to be prepared to avoid losing, deletion or archiving of audit records.
Monitoring, Review and Improve		
Monitoring the audit programme	Ineffective monitoring of audit programme outcomes	Making an audit action plan at each stage to ensure the outcomes are effectively monitored and addressed in order to meet the goal of process innovation.
Improving the audit programme	Getting information of the audit development for improving the next ones	Using information on improvement to make more efficient audits in the future (Improving)
Availability and Cooperation		
Availability and cooperation of auditee	Insufficient availability and cooperation of auditee	Allowing multiple audits to be conducted in a single visit, minimizing time and distances travelling to site, aligning audit dates with the availability of auditee's key staff
Availability of evidence to be sampled	Failure to provide samples during the audits	Evidence must be kept on file so that all members may access them, which will cut down on the amount of time needed to locate records during audits.
Content and conducting		
Answering audit criteria questions	The auditor should not be subjective when answering audit criteria questions	Try to avoid subjective evaluation, as they can lead to different results. (Facts, not opinions) Create precise, clear questions
Collecting evidence for the audit	There should be documented evidence of actions	Must upload a certified document to every affirmative answer that request it

Complexity, length and language of the audit	The simpler and more understandable the questions are the better	Make the audit as easy and understandable possible for the auditee (questions and data), separated by parts and in multilanguage
Generalization of the audit	General guide difficult to implement in industrial use cases	Specification for specific machines, data, sensors, etc. When implementing the audit
Results of the audit	Answers difficult to weigh and associate with the result characteristics	Relate questions to one of the quality characteristics, with yes/no questions or fixed answers (no free text) to be easier to weigh. Objectively quantifiable questions (not quality)

Table 4. Risks and Opportunities related to i4Q[LCP]

4.2.4 i4Q^{LCP} Audit team: Roles and responsibilities [15]

In order to establish an audit programme, various roles and responsibilities of the individuals managing the audit need to be defined. In the case of i4Q^{LCP}, these roles relate to audit processes in general (according to ISO 19011:2018 [6]) and to data quality management (according to ISO 8000-2:2020 [11] and ISO/TS 8000-150:2011 [19]. **Table 5** provides an overview of the roles, responsibilities and competencies needed in the audit team to conduct i4Q^{LCP}. **Figure 9** shows how these roles can be structured or connected.

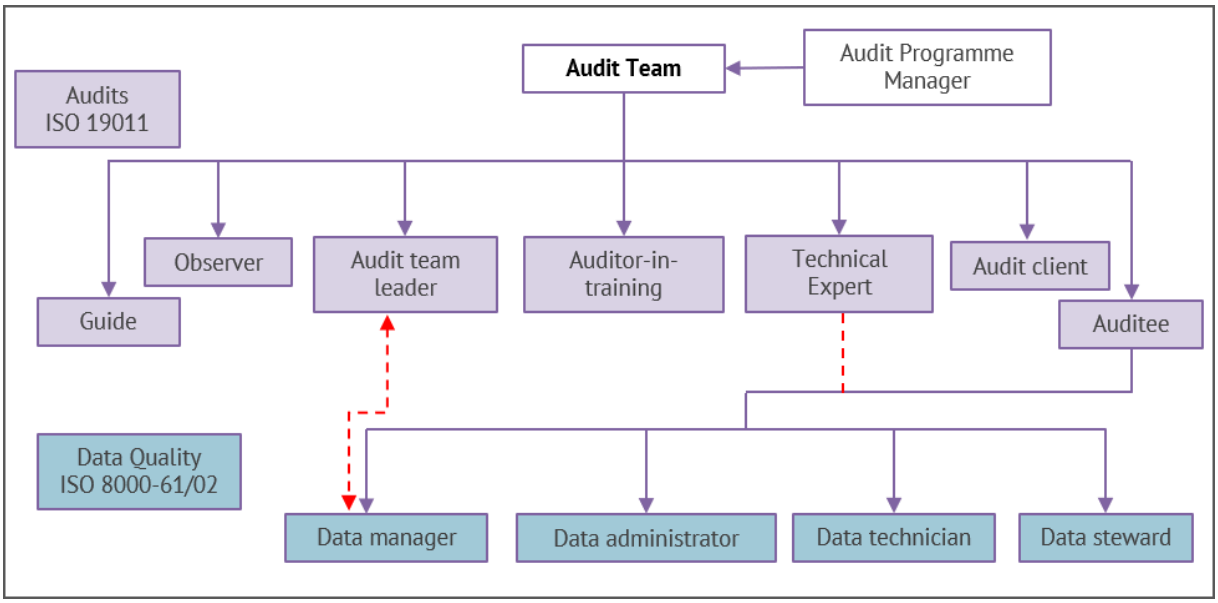

Figure 9. Audit team: Structure of roles and responsibilities in i4Q^{LCP}

Roles in i4Q^{LCP}	Responsibilities
Roles related to audits (according to DIN EN ISO 19011:2018 [6])	
Audit programme manager	• Responsible for planning, implementing, and conducting the audit programme • For i4Q^{LCP} the role could be fulfilled by a quality engineer or quality manager with experience in data management • For i4Q^{LCP} it can be the same person as audit team leader
Guide	• Assisting the auditors in collecting information (when needed) • Ensuring the rules concerning location-specific arrangements are known and followed by an audit team. • Since i4Q^{LCP} is an internal audit procedure, the audit team leader and guide are in most cases the same person
Observer	• The observer should not interact with the audit team but should neutrally observe whether all audit principles and methods are correctly applied.

[15] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

Audit team leader	• Provide significant data on audit objectives, scope, criteria, methods, and audit team • Determine appropriate statutory and regulatory requirements and other requirements relevant to the activities, processes, products, and services of the auditee • Communication channel and resolve issues with respect to the composition of the audit team and intermittently communicate the progress, any changes or any findings with the auditee or audit client • Since i4Q[LCP] is an internal audit procedure, the audit team leader and guide are in most cases the same person
Audit client (customer)	• Ensure the audit programme objectives have been achieved to direct the planning, conducting of audit and its implementation effectively.
Auditee	• An organization that is audited (here: manufacturing line, cell…)
Technical Expert	• In case of i4Q[LCP], a technical expert could be the same person as the Data manager
Auditors-in-training	• May be included in audit team under guidance of an auditor (Optional)
Roles related to data quality (according to ISO 8000-2 and ISO 8000-150)	
Data manager	• Directs a master data quality management guideline in accordance with an organization's objectives, regulates elements that affect data quality at the organizational level, and makes strategies for executing data quality processes of an organization and support activities under the role of the data administrator • Manager ensures the work practices comply with company procedures and ISO standards
Data administrator	• Controls and coordinates data technicians by defining standards for maintaining master data quality, and avoiding recurrence of data errors by evaluating reasons of errors, eliminating root causes, or developing data schema to support procedures in the function of data technician • Records for any non-conformance and performance monitoring of audit
Data technician	• The data technician creates, reads, updates, and deletes data in accordance with the data administrator's data quality management criteria, measures data quality, and corrects incorrect data as a result of the measurement • Prepare reports, retrieve required information, and dispose of outdated data of audit
Data steward	• Manages a specific set of data sources

Table 5. Audit team: Definition of roles and responsibilities in i4Q[LCP]

In general, audit programme managers should determine the scope of the audit programme according to the specific objectives and any known constraints. In addition, potential internal or external issues, risks and opportunities that may impact the audit programme should be considered and addressed. Audit team members should be selected to ensure the overall competence for all audit activities [6].

The audit team should plan, coordinate, and schedule all relevant audits. In the case of i4Q[LCP], the process can be performed in a single audit or in a series of sub-audits, suitable for the

manufacturing equipment being audited. Responsible team members should also determine and ensure the provision of all necessary resources, including the appropriate documented information and records related to the audit programme. They should also monitor, review, and improve the audit programme and finally, establish all necessary internal and external communication processes between audit team members and relevant interested parties [6]. In the case of i4Q^{LCP} interested parties could be data analysts, quality managers, shop floor workers, production engineers and managers.

To ensure that individuals managing the audit programme have all the necessary competencies, knowledge should be possessed or acquired in the following areas [6]:

- Audit principles (see Section 2.1.1), methods and processes
- Management system standards and other relevant standards and/or documents; references to relevant standards are included in i4Q^{LCP}, some explanations can be found directly in the document and IT-assistant
- In the context of i4Q^{LCP}, knowledge in the area of data quality in particular is strongly needed
- Information about the auditee, its processes and context (e.g., external/internal issues, relevant interested parties and expectations, business activities, needs, products and services)
- Applicable statutory and regulatory requirements and other requirements relevant to the business activities of the auditee.

4.2.5 Establishing extent and determining resources of i4Q^{LCP} audit programme[16]

To define the extent of the audit programme, the DLC of ISO/IEC 25024:2015 [9] (Section 2.2.1) should be considered. The audit programme of i4Q^{LCP} includes the complete DLC with all stages to ensure data quality in every process step. To address the complexity of the data quality management system, the audit programme shall be divided into individual audits, each covering one stage/phase of the DLC. The audit programme can be adapted to a machine, a cell, a complete manufacturing line or other appropriate units. The programme can then be divided into individual audits for one of the components or one machine.

When determining the resources for the audit programme, the extent of the programme should be considered. If a complete manufacturing line is to be audited, the duration may be greater than for a single machine. Risks and opportunities (Section 4.2.3) should also be considered and will impact resource requirements. In addition to time capacity and technical requirements, a key resource is the audit team itself, which should be assembled as described in Section 4.2.4. Where appropriate, the team may incur travel and accommodation costs, as in some cases not all competencies are available at each site. Technically, there is a need for knowledge in every reference to the DLC. Hardware and software for data acquisition, data storage and data processing should be available as well as data design and all documentation related to it. Auditors should be able to access all technical resources (e.g., PLC, MES) on-site or remotely.

[16] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

4.2.6 i4Q^{LCP} Audit methods [17]

Depending on the audit objective, scope and criteria, audit methods should be determined to obtain audit evidence. Audit evidence consists of records, statements of fact or other verifiable information that is relevant to the audit criteria. In order to determine conclusions, evidence should be evaluated against audit criteria (Section 2.1.2, **Figure 2**). Nonconformities and the audit findings supporting them should be documented and, if applicable, audit evidence should support conformance and effectiveness. Another aspect to consider is the sampling technique, which is used when records are too numerous or too dispersed to warrant examination of all elements of the population and it is not feasible or cost effective to review all information during the audit. Two different kinds of sampling techniques are used [6]):

- **Judgement-based sampling**: The skills and expertise of the audit team, audit experience within the audit scope, previously identified significant risks and opportunities for improvement, and the complexity of the requirements to achieve the audit objectives are important factors to consider when performing judgment-based sampling. A disadvantage of these sampling techniques is that they do not allow for a statistical assessment of the impact of uncertainties in audit findings and conclusions.
- **Statistical sampling**: The sampling strategy should be based on probability theory, i.e., the audit objectives and what is known about the general characteristics of the population from which the samples are drawn. Attribute-based sampling is used when there are only two possible sample outcomes for each sample and variable-based sampling is used when the sample outcomes are in a continuous range.

Table 6 shows what types of on-site, remote, or a combination of these two audit methods are possible.

Type of Interaction	Location of Auditor	
	On-Site	**Remote**
Human Interaction	- Conducting interviews of auditee - Completing checklists and questionnaires with auditee participation related to data quality - Conducting document review with auditee participation according to external/internal standards, conventions, and regulations	Via interactive communication means: - Conducting interviews - Observing work performed with remote guide - Completing checklists and questionnaires - Conducting document review with auditee participation according to external/internal standards, conventions, and regulations

[17] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

No human interaction	- Conducting document review (e.g., records, data analysis, calibration reports) according to external/internal standards, conventions, and regulations - Observing and monitoring the data processes - Completing checklists of audit criteria	- Conducting document review (e.g., records, data analysis, calibration reports) according to external/internal standards, conventions, and regulations - Observing and monitoring the data processes - Completing checklists of audit criteria

Table 6. i4Q^{LCP} Audit methods [6]

Table 7 presents a complete description of the process steps of collecting and verifying information in i4Q^{LCP} (According to **Figure 3**).

Process steps of collecting and verifying information (cf. DIN EN ISO 19011:2018)		Implementation in i4Q^{LCP}
1	Source of information	Interviews with e.g., process owner
		Documents (upload or link)
		Data sets
2	Collecting by means of appropriate sampling	Collect correct samplings related to the specific machine, cell or manufacturing line
		Use appropriate sampling techniques
		Evaluate representative and reliable data base
3	Audit evidence (information that be verified, if verification is not possible the auditor should judge if evidence is given)	Calibration Protocols
		Measurement Protocols
		Testing protocols
		Data streams
		Data files (e.g., .csv files), provided as an upload or link
		Record the audit evidence (upload or link)
4	Evaluating against audit criteria	Evaluation is done automatically in IT assistant
		Sufficient number of relevant questions related to audit criteria to be answered in the IT assistant
		The procedure cannot be changed in the order of the steps
		Mandatory questions cannot be skipped
5	Audit findings (indicate conformity or non-conformity)	Evaluation of maturity level of data quality in the audited manufacturing line or cell
		Opportunities and recommendations to improve the data quality
		Inadequate procedure steps for obtaining the certification
		Risk identification in data life cycle
		Identification of good practices
6	Reviewing	Review complete process and results
7	Audit conclusions	Data Quality Certification is decided by the auditor
		Recommendations for further actions and audits
		Audit conclusions are shared with all relevant stakeholders
		Audit report is provided by digital tool in PDF format

Table 7. Process steps of collecting and verifying information in i4Q^{LCP}

4.2.7 Documentation, monitoring, and improvement of i4Q^{LCP} audit programme[18]

According to ISO 19011:2018 [6], audit records should be created during the implementation and conduct of the i4Q^{LCP} audit programme to provide information on the achievement of objectives. These records should be managed and maintained by the auditors and may document the following information related to the audit programme and/or each audit:

- Risks and opportunities, external and internal issues, and feedback of audit programme,
- Audit plans, audit criteria report, audit findings and evidence report, non-conformity reports, corrective action reports and follow-up reports
- Audit team performance evaluation, criteria for selection of audit team, maintenance, and improvement of competence

The documented information for the audit can include but is not limited to: Physical or digital checklists, audit sampling details and audio-visual information. The format of the records also depends on the audit methods used. Especially the audit criteria report should be supported and/or automatically created by the i4Q^{LCP} IT-tool described in Section 5.

These records support the management of the audit programme results, i.e., evaluating the achievement of objectives, creating, reviewing and approval of audit reports, as well as distributing them to interested parties and reviewing the effectiveness of actions and determining further audits if needed. Consideration should also be given to passing on good audit results to other parts of the company. So, if high data quality has been achieved in one part of production, other units should be able to follow ([6], p. 38). Additionally, the records generated should support the monitoring of the audit programme which includes the evaluation of the following aspects ([6], p. 40):

- Whether audits are being conducted as per schedule and objectives are being achieved
- Performance of all audit team members and their ability to implement the audit plan
- Feedback of the audit programme
- Whether the audit criteria document is being followed
- Whether the documented information in the audit process is completed

Auditors should document and monitor whether changes to the audit programme are required. When reviewing the audit programme or any audit, specific risks and opportunities that occur should be included in the list in Section 4.2.3. In addition, audit records and methods should be reviewed and adjusted as necessary to improve the specific audit process ([6], p. 41). As part of i4Q project, the i4Q^{LCP} audit programme will also be revised and improved as necessary during the validation phase in WP6.

[18] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

4.2.8 Overview of implementing i4Q^{LCP} audit programme [19]

To implement the audit programme, the steps described in ISO 19011:2018 ([6], p. 32) and shown in **Figure 10** should be performed by the audit team members. All necessary and specific information about implementing i4Q^{LCP} can be found in the mentioned sections of this guideline, including:

- Defining the objectives, scope and criteria for an individual audit (see sections 4.2.1, 4.2.2 and 4.3)
- Selecting and determining audit methods (see section 4.2.6)
- Selecting audit team members (see section 4.2.4)
- Assigning responsibility for an individual audit to the audit team leader (see section 4.2.4)
- Managing risks and opportunities related to the audit programme (see section 4.2.3)
- Management, monitoring, and improvement of audit programme results and records (see section 4.2.8)

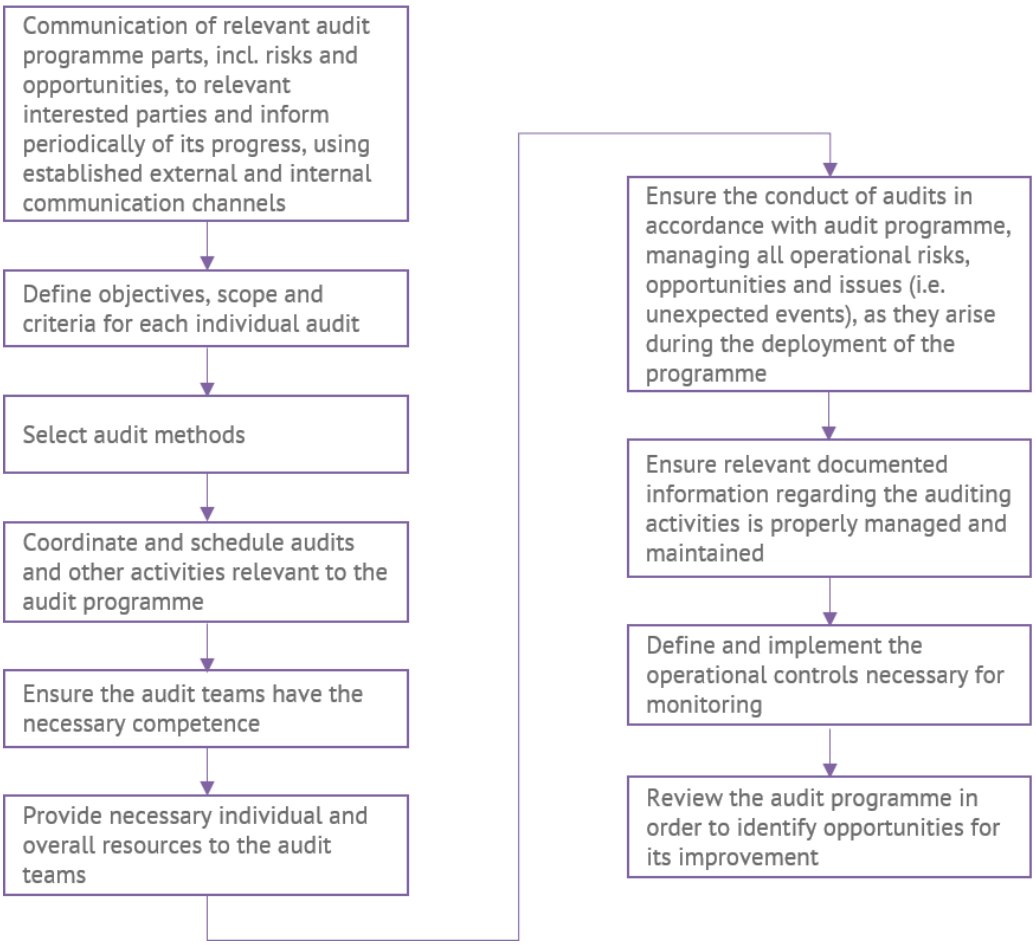

Figure 10. Process of implementing the audit programme according to ISO 19011:2018 [6]

[19] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

4.2.9 Procedure of conducting i4Q^{LCP} audit

To conduct a specific audit of the i4Q^{LCP} audit programme according to the aspects described in Section 4.2.1 to 4.2.8, **Figure 11** shows a general overview of a process chart with all relevant steps that are necessary to an i4Q^{LCP} audit. This is a detailed description of the last process step in **Figure 8**. For further information, ISO 19011:2018 provides a detailed guide ([6], p. 42) to the individual steps and aspects to be considered. As highlighted in **Figure 11**, the i4Q^{LCP} IT-assistant tool (Audit Advisor) supports several steps in the core stages of audit conduct. The IT-assistant can be used for preparation and conducting the audit by answering the audit criteria questions and use the report generated as a guide for fulfilling the tasks suggested by IT-assistant. Additionally, this report can be used as basis for creating final audit report by the auditor and keep it as evidence for the audit findings.

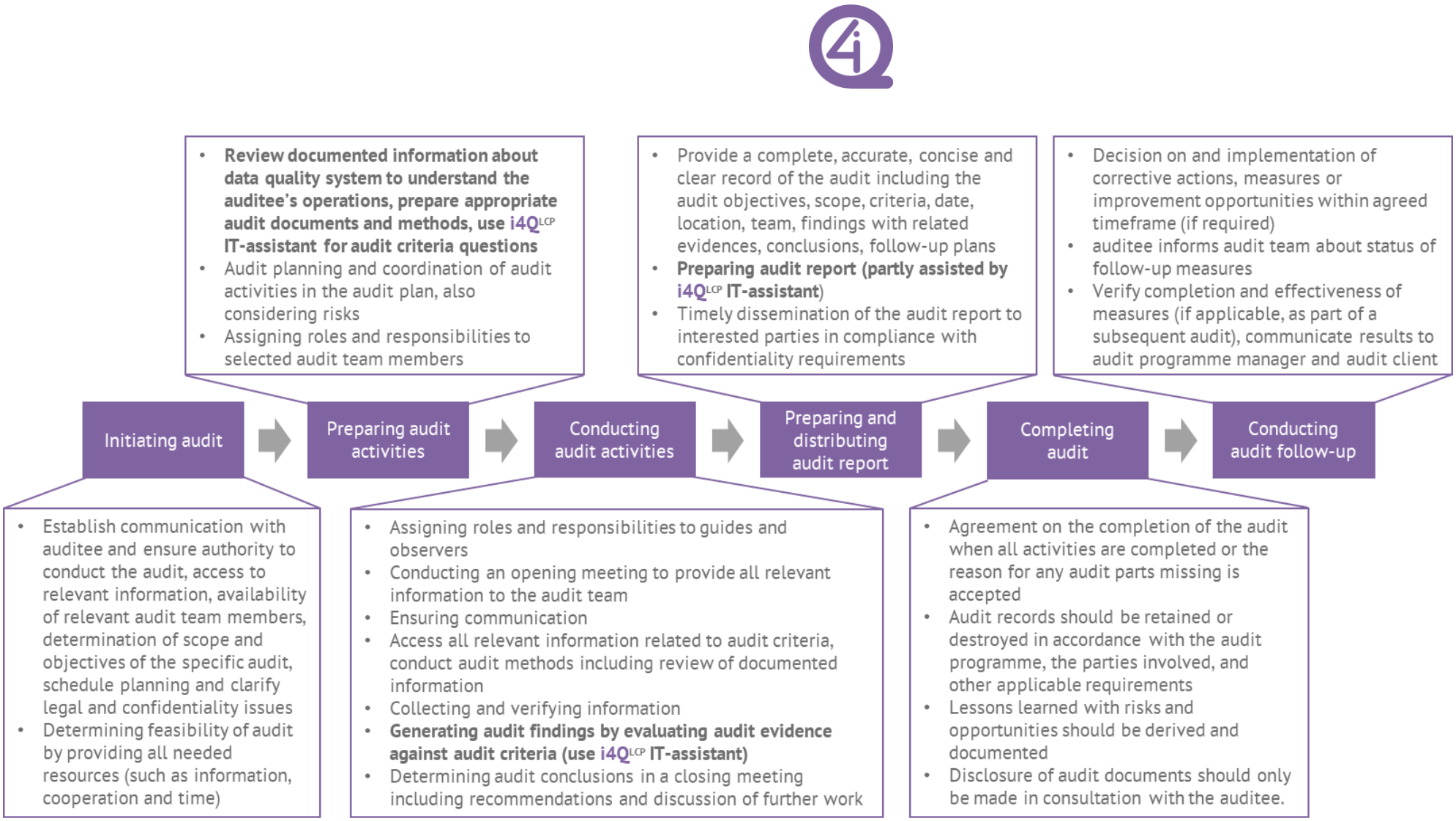

Figure 11. Conducting an i4Q LCP audit according to ISO 19011:2018 [6]

4.3 i4Q^{LCP} Audit Criteria

This section describes the audit criteria that are used in i4Q^{LCP} Manufacturing Line Data Certification Procedure. These criteria consist of requirements whose fulfilment must be confirmed by objective evidence in order to enable the auditee to pass the audit.

4.3.1 Audit Criteria Structure

The audit criteria are composed of questions based on the data quality model and data quality measures as defined in (ISO/IEC 25012:2008 [8], ISO/IEC 15939:2017 [5] and ISO/IEC 25024:2008 [9]. The data quality measures for measuring the data quality in terms of quality characteristics are defined as per (ISO/IEC 25012:2008 [8]).

The audit criteria used in i4Q^{LCP} Manufacturing Line Data Certification Procedure is structured into a number of data life cycle (DLC) phases such as Data collection, External data acquisition, Data integration, Data processing, Presentation etc. A separate phase "Measurement System" is developed as per ISO/IEC 15939:2017 [5] which will be executed as a pre-requisite for the audit criteria before the execution of any DLC phases. **Figure 12** shows the overall Audit structure.

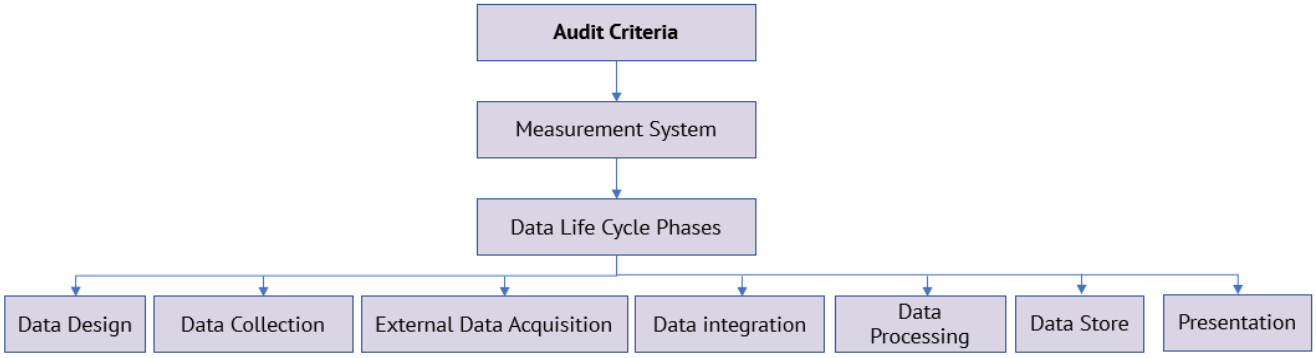

Figure 12. Audit Structure

The sequence of execution for each DLC phase is described in Section 4.3.2. Each of the DLC phases is associated with Target Entities as shown in **Figure 13**.

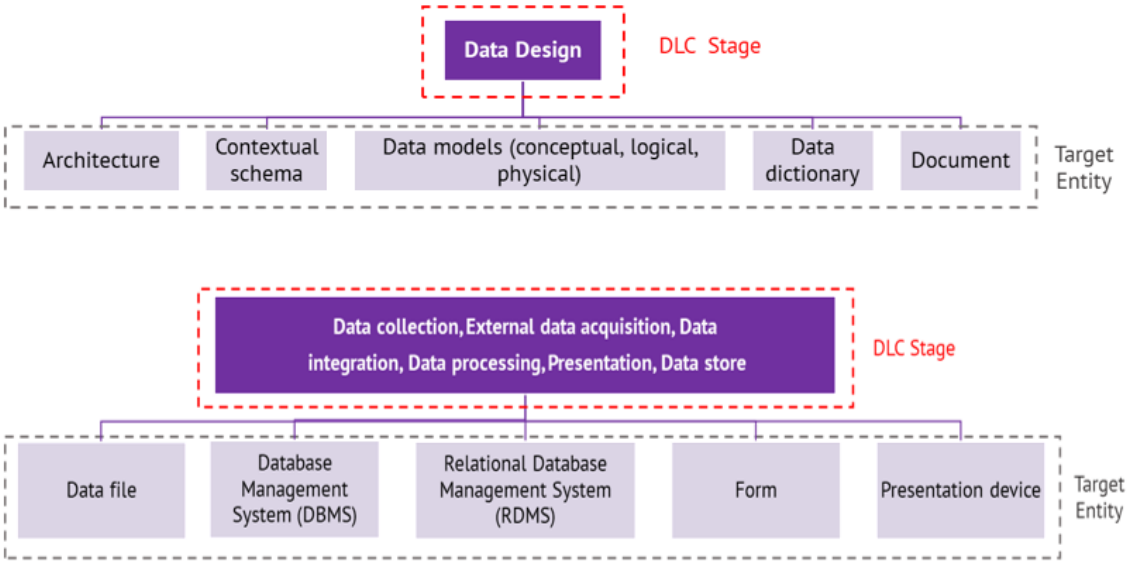

Figure 13. Target Entities associated with each DLC phase

Figure 14 depicts properties associated with the Target Entities**.**

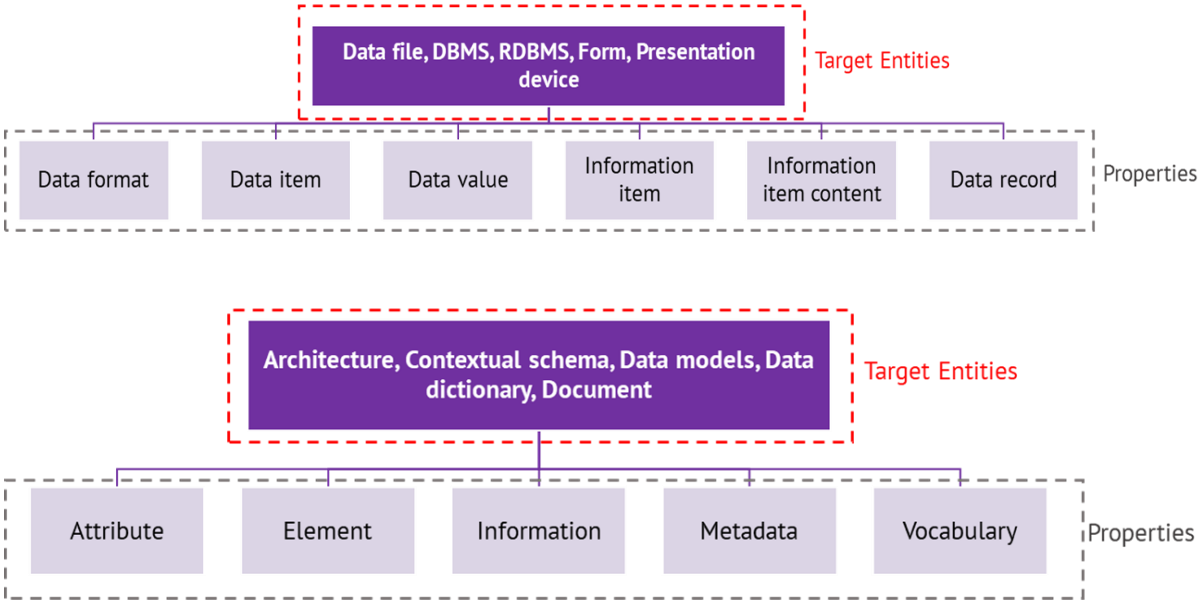

Figure 14. Target Entities with properties

Section 4.3.3 describes the target entities and their associated properties in detail.

4.3.2 Data Life Cycle

The data life cycle is the series of steps that a specific piece of data takes from the time it is created or captured until it is eventually archived and/or deleted at the end of its useful life [29]. Various Target Entities are connected to each step of the DLC and each target entities are associated with properties. The stages that make up the data life cycle are as follows: data design,

external data acquisition, data collection, data integration, data processing, data presentation, data store, and other use. **Figure 15** depicts the sequence of DLC stages.

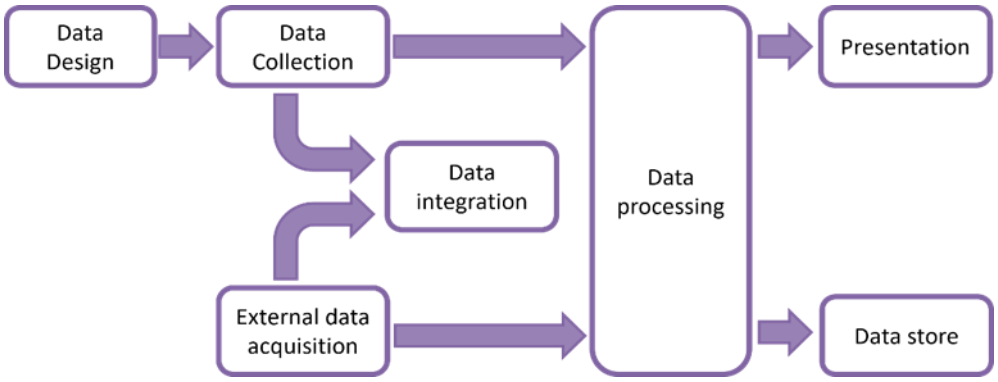

Figure 15. Stages of Data Life Cycle (adapted from ISO/IEC 25024:2015)

Data Design: This stage the organisation describes the process by which data type is created and values enter system. The data items and data models are defined as per the organisation needs and usage. Designing of data comprises of architecture, conceptual data model, logical and physical data model.

Example: A set of numbers, symbols or alphabets in defined format which create a data item.

Data Collection: In this stage organisation gets information all from the periodically data generated. The data team is responsible for determining what data needs to be collected and the most effective ways to do so, as well as what information is unneeded or irrelevant to the business needs. It is crucial to remember that many businesses use a broad strategy for data collection, gathering as much information as they can from each engagement. However, the organisation using the data is responsible for the data quality.

Example: Data can be gathered in many different methods, such as through forms, surveys, interviews, direct observation, electronic devices, and sensors and more.

External Data Acquisition: This stage is associated with various target entities viz. data file, DBMS, RDMS, form and presentation device. When obtaining data, several factors need to be taken into account. In this phase, an organization collects the data, which another organisation then acquires the whole for further business use. As soon as the data are acquired or received, they need to be examined to make sure they adhere to standards and may be approved for their intended purpose. Business requirements, data standards, accuracy standards, cost, time limitations, and format considerations are commonly applied when acquiring data.

Example: acquiring and capturing data created outside of an organization by means of physical documents, electronic devices, internet of things (IoT) devices, forms, surveys, and more.

Data Integration: This stage refers to processes that attempt to glean meaningful insights from raw data [29]. It emphasises on the degree to which the data is precise according to the specification. Target entities and properties like data file and data item respectively focuses on the completeness, accurate and consistency of the data. Processing raw data and turning it into information can be exceedingly labour-intensive. A scientific procedure is developed to integrate the raw data and to record the data inputs. Because it is crucial to record the process of how the data are examined and converted after collection, workflow capture is significant within the data

lifecycle. The analyst should keep track of the entire process since doing so improves transparency and reproducibility and makes it simple for others to monitor the development of the data analysis in the future.

Example: Integrating sensor data from two different databases with different data models to serve the same frontend.

Data Processing: As part of the data processing phase, data is frequently reformatted, summarized, subset, standardized, and enriched. Data wrangling is the process of cleaning and transforming a data set from its unprocessed state into one that is more useable and accessible. This is sometimes referred to as data clean-up, data cleansing, or data munging [29]. At times the sensitive data is translated into another form of code to protect it from privacy concerns is called data encryption which is as well a part of data processing.

Example: Calculating statistics for a time series (e.g., mean, maximum and minimum values).

Presentation: The process of visualizing data is known as data presentation, and it is normally carried out using one or more visualization tools. By using data visualization, one can more easily explain analysis to people inside and outside of an organization. The format that visualization takes relies on the data being used and the narrative that is being conveyed [29]. This phase of data life cycle provides the opportunity to make sense of your analysis and visualization.

Example: Making data discoverable by publishing metadata in data catalogues, portals, or blogs. Sometimes specific keywords and tags are used for presentation.

Data Store: Data stored in basic backbone support for data processing. In this case, data is important for various purposes for industries to meet objectives and operations. However, data cannot reproduce or recycle it, somewhere it needs to be re-using the same for another purpose [25]. It is critical to keep information about each item in the records, especially data origin. Data store involves protecting data from accidental data loss, corruption, and unauthorize access. Data storage and the creation of backups of acquired data are both crucial components of data management. Backups protect information against human error, hardware malfunction, virus assaults, power outages, and natural disasters. If these errors do occur, backups can help organizations save time and money.

Example: Storing data in cloud systems physical secured computer.

4.3.3 Terms and definitions of Target Entity and Properties

Fundamental thing of relevance to the user, about which information is kept, and need to be measured. Target entities are input to information product and work product. Examples of target entities are architecture, contextual schema, conceptual and logical and physical data models, data dictionary, document, data file, database management, relational database management system, form, and presentation device are described in **Table 8**.

Target Entity	Description	Examples
Architecture	Fundamental concepts or properties of a system in its environment embodied in its	Concept how e.g., sensors in machine are connected to data base, which is connected to hard-

Target Entity	Description	Examples
	elements, relationships, and in the principles of its design and evolution. [9]	and software resources for analysis and how data is generated, handled, transformed and stored.
Contextual schema	Formal description of the boundary of the context of use where data models are applied. It includes a holistic vision of a (system) context of the architecture. [9]	It describes what information is needed to fulfil a specific use case in a generic way e.g., To ensure motor is not overheating temperature data is needed.
Data models	Graphical and textual representation of analysis that identifies the data needed by an organization to achieve its mission, functions, goals, objectives, and strategies and to manage and rate the organization. A data model identifies the entities, domains (attributes), and relationships (associations) with other data and provides the conceptual view of the data and the relationships among data. [9]	Data such as RPM of tool, force exerted and encountered by tool etc. during the facing operation simulation is used to find/predict tool life and tool wear.
Data dictionary	Collection of information about data such as name, description, creator, owner, provenance, translation in different languages, and usage. [9]	
Document	Records any aspect of project design, sampling, data collection, cleaning and analysis that provides information or evidence[20]	Job sheet used by an operator in a CNC Machine to document key tasks performed and used also in quality audits.
Data file	Set of related data records treated as a unit. [9]	Number of job sheets for a particular type of product produced by a machine over a particular period.

[20]https://dimewiki.worldbank.org/Data_Documentation#:~:text=Data%20documentation%20is%20the%20process,course%20of%20a%20research%20project

Target Entity	Description	Examples
Database Management System	It is a structured collection of data and software designed to store, retrieve, define, and manage data in a database. [9]	MySQL, PostgreSQL, Microsoft Access, SQL Server, FileMaker, Oracle
Relational Database Management System	Management system for relational database i.e. system for storing, managing, querying, and retrieving data from a relational database. [9]	Microsoft SQL Server, Oracle Database, MySQL and IBM DB2
Form	A form is a user interface element that allows you to enter and submit data on a software application i.e. module or formulary to collect data [9]	In the drilling function present on the CNC control panel, the user has to fill various parameters such as drilling location, dept of drill, coolant activation etc in the blanks.
Presentation device	Device used to present data to the intended user of a system. module or formulary to collect data. [9]	CNC Control Panel, Dashboard, handheld device, etc.

Table 8. Terms and definitions of Target Entity

The properties of the database provide detailed information about the database. Target entities are precisely defined by properties. Examples of properties are attribute, element, information, metadata, vocabulary, data format, data item, data value, information item, information item content, and data record [27] are described in **Table 9**.

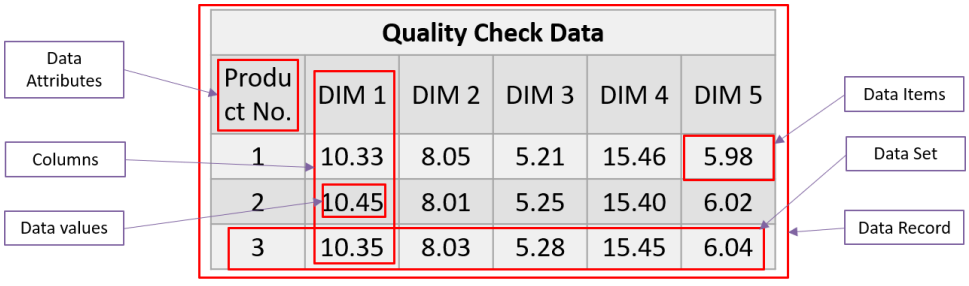

Figure 16. Example of use of Audit Criteria Vocabulary

Properties	Description	Examples
Attribute	Inherent property or characteristic of a target entity that can be distinguished	In above Product No., DIM 1, DIM 2 are examples of data attributes.

Properties	Description	Examples
	quantitatively or qualitatively by human or automated means. [9]	
Element	Smaller part of an architecture. [9]	Data pipeline, Sensor, Data Source
Information	In information processing, knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context have a particular meaning. [9]	Sets data into context
Metadata	Data that describe other data. [9]	File Name, Properties, Created by, Modified By, File Size, File Type etc of an MS Excel file.
Vocabulary	Collection of information related to a specific subset of terms related to a specific domain. [9]	
Data format	Arrangement of data for storage or display	Quality data stored in tabular form.
Data item	Smallest identifiable unit of data within a certain context for which the definition, identification, permissible values, and other information is specified by means of a set of properties. [9]	In quality data, the cell with the value 5.98 of Product No. 1.
Data value	Content of data item. [9]	In quality data, the data value of Product No. 2's first dimension (DIM1) is 10.45.
Information item	Separately identifiable body of information that is produced, stored, and delivered for human use. [9]	Abstract classifier representing information, which is transferred within a system, e.g., digital Order is send by the customer over the system to the shop. It contains information about quantity, price, and delivery date.
Information item content	Information included in an information item, associated with a system, product, or service to satisfy a requirement, or need. [9]	

Properties	Description	Examples
Data record	Set of related data items treated as a unit. [9]	Quality Check data table including all data is Data record.

Table 9. Terms and definitions of Properties

4.3.4 Audit Criteria Questions

For each quality characteristic assigned to a target entity and its particular property, a number of consecutive questions have been formulated. If the auditee's answer to a question is "Yes", the auditee proceeds to the next questions until the last question for the particular quality characteristic is reached. If the answer is "No" for any question, the audit is put on hold and a corrective task is recommended to address the identified data quality deficiencies.

Table 10 explains the columns of the MS-Excel Audit Questions. These audit criteria questions will be integrated with the IT tool as described in Section 5.2. **Table 11** shows the exemplary audit criteria questions for DLC data collection. Further examples can be found in **Appendix 8.4.**

Column Title	Column Description
Target Entity	It is fundamental of relevance to the user, about which information is kept, and need to be measured (ISO/IEC 25021:2012, 4.17). For example: architecture, contextual schema, conceptual and logical and physical data models, data dictionary, document, data file, database management, relational database management system, form and presentation device.
Properties	It refers to the particular property of the target entity.
Description	It describes the property assigned to the target entity.
Data Quality Characteristics	Refers to the quality characteristics relevant to the particular target entity.
ID	Identification code (or identifier) of a data Quality Measure. ([9])
Description	It defines the quality characteristics.
Questions for Guideline	These are the questions listed for the particular target entity in the data life cycle phase.
Answer	Answer to the question which are either “Yes” or “No”.
Task	Recommended Task for the auditee to be undertaken when the answer to an audit question is “No”.

Table 10. MS-Excel Audit Questions –Definition of Columns

Target Entity (ISO/IEC 25024:2015(E))	Properties (ISO/IEC 25024:2015(E))	Description (ISO/IEC 25024:2015(E))	Data Quality Characteristics (ISO/IEC/IEEE 25012:2008).	ID	Description (ISO/IEC/IEEE 25012:2008)	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task
Data File	Data Format	arrangement of data (4.5) for storage or display	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes				
							No	Data values shall be defined and documented according to external/internal standards, conventions, and regulations.		No	Define, establish and document a process for use of internal/external standards, conventions, and regulations for data arrangement.		No	Provide related persons access to documentation of all used internal/external standards, conventions or regulations.			
Data File	Data Format	arrangement of data (4.5) for storage or display	Precision	Pre-D-1	Degree to which data format keep precision according to the specification	Do specifications on precision of data format (e.g. data type in a table or length of data items) exist in your company?	Yes		Is the specification on precision of data format similar to the internal data collection standard?	Yes		Are processes defined and established ensuring that precision of data format specifications are met?	Yes		Are the specifications for precision of data format communicated to all relevant persons?	Yes	
							No	Specifications on data format shall be defined and documented.		No	External data acquisition specification on precision of data format shall be similar to the internal data collection standard		No	Processes ensuring that precision of data format specifications are met shall be defined, documented and established.		No	Specifications on precision of data format shall be communicated to all relevant persons.
Data File	Data Item	smallest identifiable unit of data (4.5) within a certain context for which the definition, identification, permissible values, and other information (4.21) is specified by means of a set of properties	Accuracy	Acc-I-1	Ratio of closeness of the data values to a set of values defined in a domain	Does your company have a defined process to validate syntactic accuracy for external data values?	Yes		Is the documented process available to relevant personnel?	Yes		Is the documented process available to relevant personnel?	Yes				
							No	A process to validate syntactic accuracy of external data values shall be defined, implemented and documented		No	Access shall be provided to the relevant personnel.		No	Access shall be provided to the relevant personnel.			

Table 11. Examples of Audit Criteria Questions for DLC- Data Collection

5. i4Q^{LCP} Digital Support Tool

This section uses and updates information from i4Q's deliverable D5.6. It describes how software can support the audit and certification procedure described above. In i4Q, this support focuses on preparing and conducting an audit. It assumes the audit criteria, which consist of specific requirements for manufacturing line data, represented by questions with binary answers:

- Yes, means a requirement is fulfilled. Requires evidence to become a conformity.
- No, means a requirement is not fulfilled. Requires a task to meet the specific requirement.

The preparation support helps auditees to identify relevant questions and document evidence that the organization meets the requirements above. This step occurs before the actual audit. If requirements are not met, the auditee needs to apply measures. While conducting the audit, the auditee can use the software to present documented evidence to accelerate the audit.

The assistance focuses on the following aspects:

- Identify relevant requirements based on the organization's context
- Guide the user through the relevant requirements
- Store the textual evidence for fulfilled requirements
- Indicate tasks needed to fulfil requirements

i4Q planned two tools to realize the assistance: First, a **visual digital checklist** provides basic assistance for the aspects above. It represents an easy-to-implement and customize assistance tool. Second, a **chatbot** represents the assistant in the narrow sense providing intuitive access to information. It can understand natural language and context to help users without inducing cognitive overload. The latter can easily occur when an IT-tool provides too much or too complex information at once. A second advantage is that organizations can extend the chatbot to provide other useful information relevant to a manufacturing line data audit. Its main disadvantage is that the tool is more complex and, therefore, costlier to develop, deploy, and maintain compared to the visual digital checklist.

However, according to experiments with the first version of visual checklists [32], the effort of building and maintaining a visual digital checklist is very high. The selected tool "MS Forms" limits forms to 200 questions. This means that we cannot implement most tabs in the Excel via one form. Splitting them into multiple forms increases the maintenance substantially. Also, with ~1000 questions total, the checklists are hard to manage with MS Forms and most likely with similar tools as well. Therefore, this deliverable will focus on the IT-based assistant design and implementation.

This section summarizes the design and implementation of the i4Q^{LCP} IT-based assistance chatbot (Audit Advisor) containing intuitive dialogs and an audit process evaluation feature.

5.1 Goals

The main goal of the IT-based assistant for i4Q^{LCP} is to help the auditee to prepare and conduct the audit. **Figure 17** illustrates the architecture for the audit advisor tool.

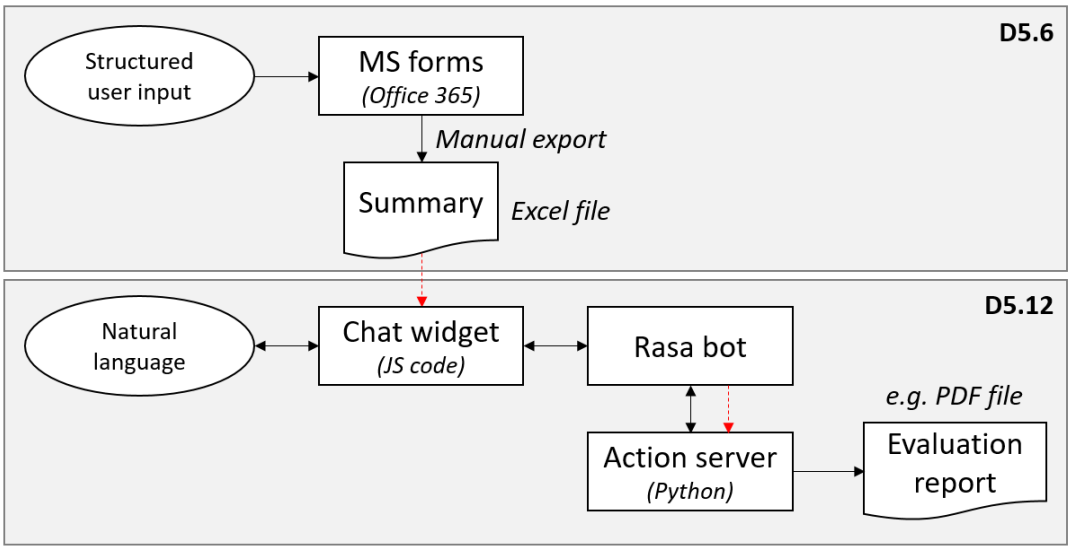

Figure 17. Architecture of the i4Q^{LCP} AI-based assistance tool "Audit Advisor"

The i4Q^{LCP} assistance is a chatbot realized with the Rasa framework. The bot processes natural language and decides on how to respond accordingly. Responses use the contents of the questionnaire above but reorganized as a knowledge graph. A special response is the evaluation report that summarizes the user's answers and tasks to meet all requirements. The advisor shares its infrastructure with the assistant i4Q^{QE} developed in T3.1.

The basic requirements identified for the i4Q^{LCP} IT assistance tool are listed below:

- Provide user-friendly interface
- Provide web-based application
- Provide a storage system
- Provide possibility to upload files and add links
- Support creation of audit report
- Create PDF document serving as an internal certificate
- Optional: User identification with password
- Optional: Project identification

5.2 Implementation specification

The specification covers several parts of the final i4Q^{LCP} chatbot including i4Q^{LCP} audit criteria questionnaires and the audit advisor software.

5.2.1 Audit criteria questionnaires

The questions based on the i4Q^{LCP} audit criteria (Section 4.3) represent the requirements on data quality an organization has to fulfil to pass the i4Q^{LCP} audit. A Data Quality Life Cycle Model (Section 4.3.2) organizes the questions along several life cycle steps. It reduces the questionnaire's complexity by breaking it down into smaller sections. Possible answers are "yes" and "no" and users can only choose one. Some questions ask for evidence in a free text field (e.g., a weblink or a short description). Each question belongs to a Data Quality Characteristic (Section 2.2.2). The

questions, answers and characteristics are linked to the measurement system defined in Section 3.2. The evaluation criteria consider the weight each characteristic has in the overall scope.

5.2.2 Audit advisor software

This software covers natural language processing and the generation of the evaluation report. Since the i4Q[LCP] auditee is mainly interested in the report, it is focused on this specification. An audit is either passed or failed, therefore, the **evaluation report** must conclude with one of these values. In addition, it should reveal the **degree of fulfilment** per quality characteristic and lifecycle stage. The audit programme manager can use this insight to identify weak spots that could be focused, for instance, in future internal audit programmes. Finally, the report includes the list of **all tasks as recommendations**.

Anytime during the audit procedure, the current answers will be saved.

5.3 Technologies

The i4Q[LCP] assistant is a standalone solution realized with Rasa. A knowledge graph stored in a Neo4J database persists the questionnaire and the audit preparation projects. Each project contains answers and the suggested tasks. A GraphQL interface allows Rasa to access the information in the graph quickly and flexibly. A simple website delivers the Rasa chatbot through a widget.

Figure 18 illustrates the knowledge graph.

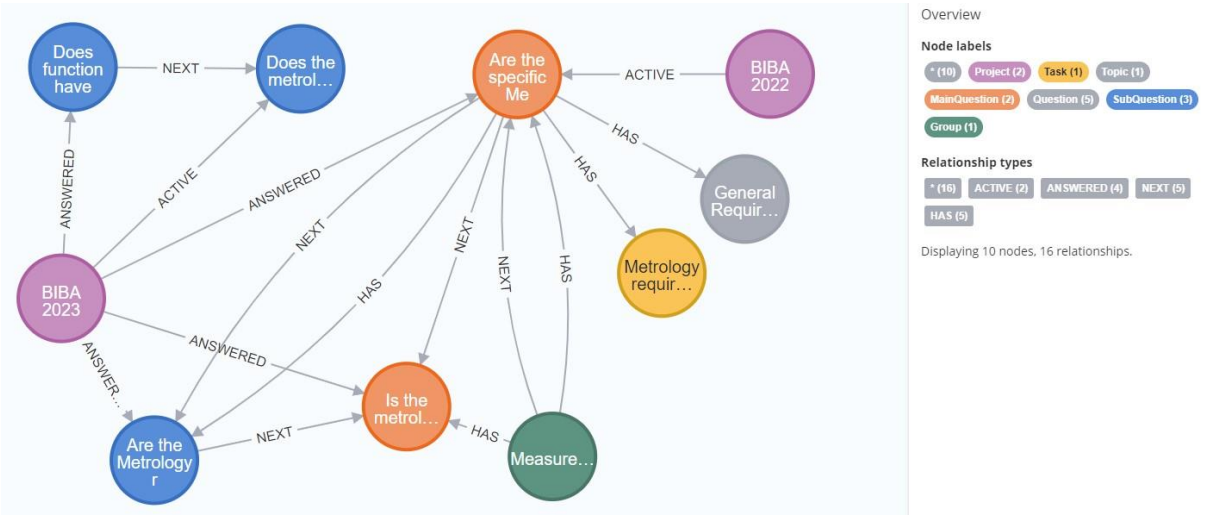

Figure 18. Knowledge graph for the audit advisor (sample graph)

5.4 User story and dialogs

This section outlines the main user story for the i4Q[LCP] IT-based assistant. Besides this main story, the Rasa chatbot has to consider various cases where users divert from the main story.

5.4.1 Basic interactions

The following interactions are basic because they result from legal obligations (e.g., emerging AI Law) and best practices of building chatbots. The entire conversation is a series of turns where the user and the bot exchange their messages.

All interactions begin by opening the chat widget on a website (e.g., the QualiExplore website). The bot will initiate the conversation as illustrated in **Figure 19**.

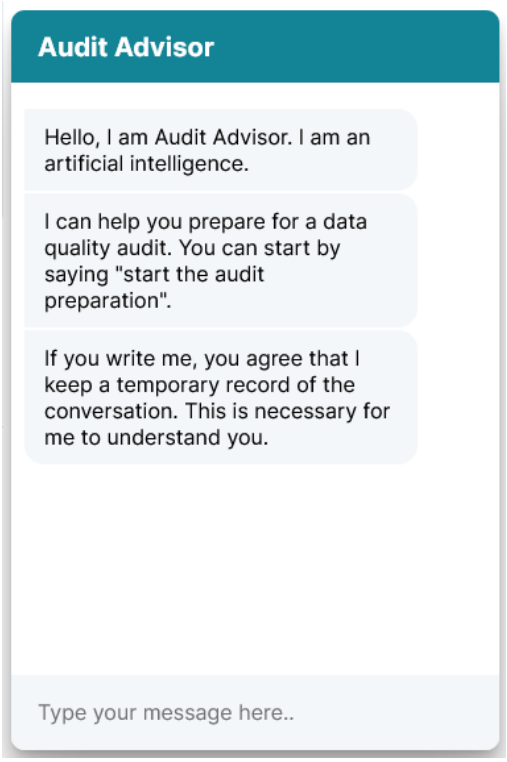

Figure 19. Bot reveals itself and opens the conversation

The bot revealing itself is an obligation resulting from the emerging EU AI Act (Section 5.2.4.). Typically, chatbots also inform the user about their purpose in the first conversation turn.

Chatbots can create and record personal information about users, e.g., log their IP address, maintain a conversation history, or store their name to personalize the interaction. Therefore, chatbots are subject to the General Data Protection Regulation (GDPR) and its legal obligations. It informs the user in the first turn about the fact that it may collect personal information.

Any further turns depend on the user – users familiar with the bot will likely ask more specific and directed questions while novices will likely be more explorative.

A typical basic interaction is asking for the bot's capabilities to understand the features. Capabilities for the assistant are illustrated in **Figure 20**.

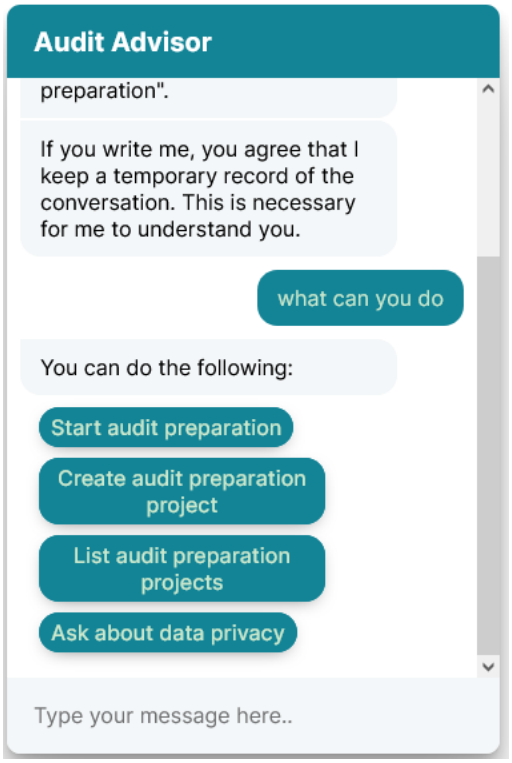

Figure 20. User asking for the assistant's capabilities

Besides the initial GDPR-related information, the chatbot has features to answer questions related to the GDPR. **Figure 21** summarizes some questions.

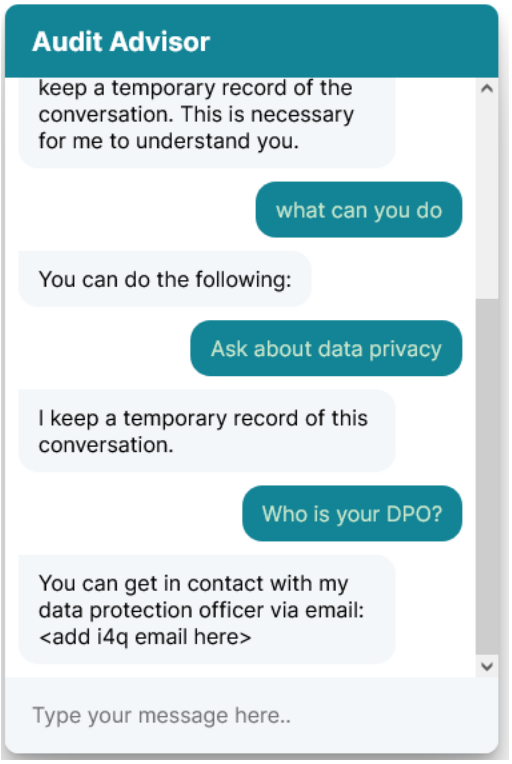

Figure 21. User asking for information about data privacy

5.4.2 Audit preparation project interactions

The following interactions are necessary so that users can create projects. Projects contain the user's answers to audit questions and the tasks. Users create a project by giving the chatbot the project's name as illustrated in **Figure 22**.

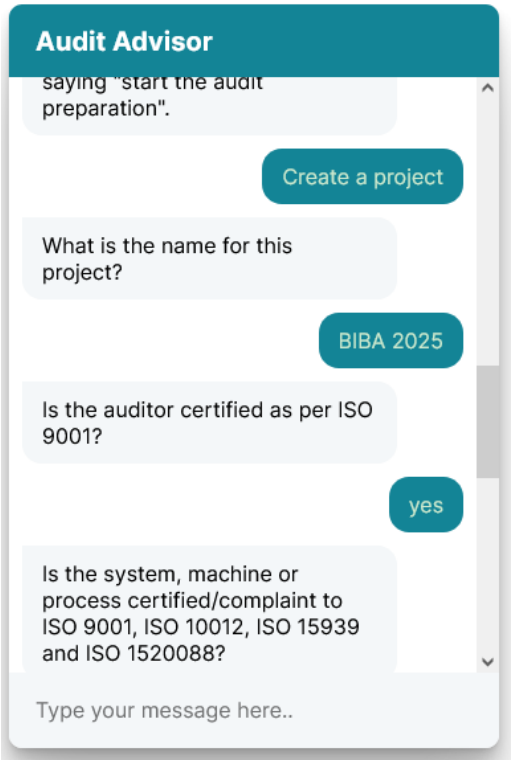

Figure 22. User creates project

They can continue an existing project as well, as indicated in **Figure 23**.

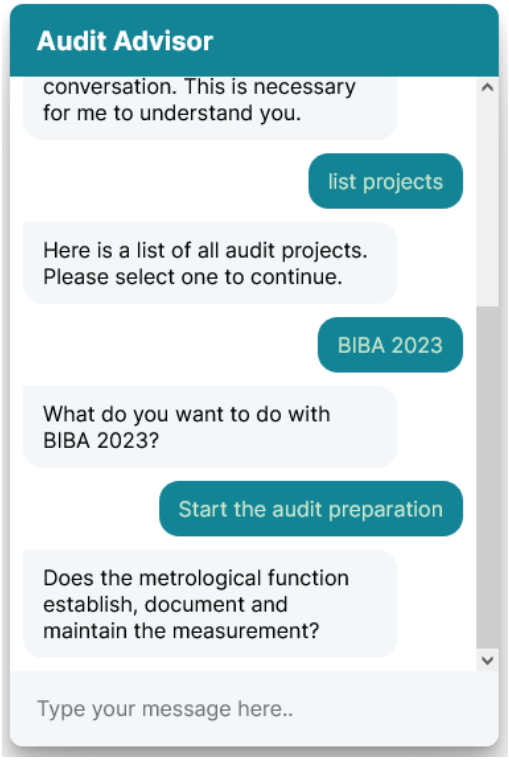

Figure 23. User continues with a project

5.4.3 Questionnaire interactions

These interactions focus on starting, answering, and completing the questionnaire - i.e., prepare a data quality audit. First, the user has to start the questionnaire either by creating a project or continuing an existing one. **Figure 24** presents how a user starts the questionnaire of a recently created project.

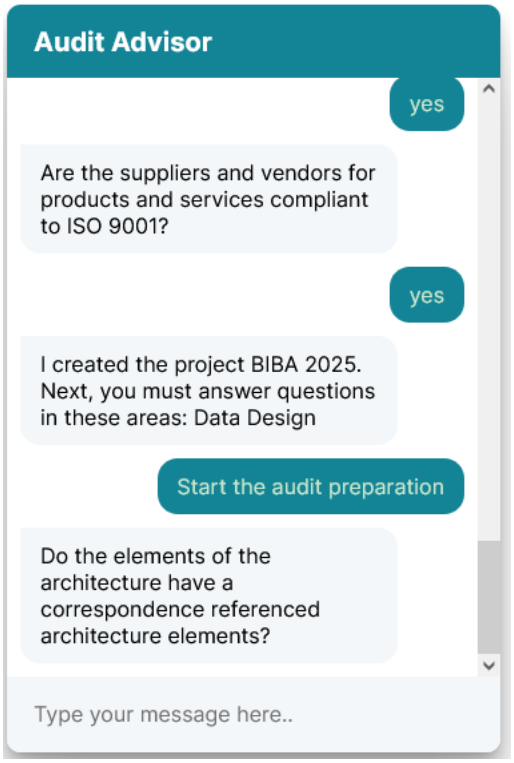

Figure 24. User starts the questionnaire

The bot provides a PDF file once the user completed the questionnaire. This file is the audit preparation report containing all questions, answers, and the tasks. **Figure 25** shows the bot's response.

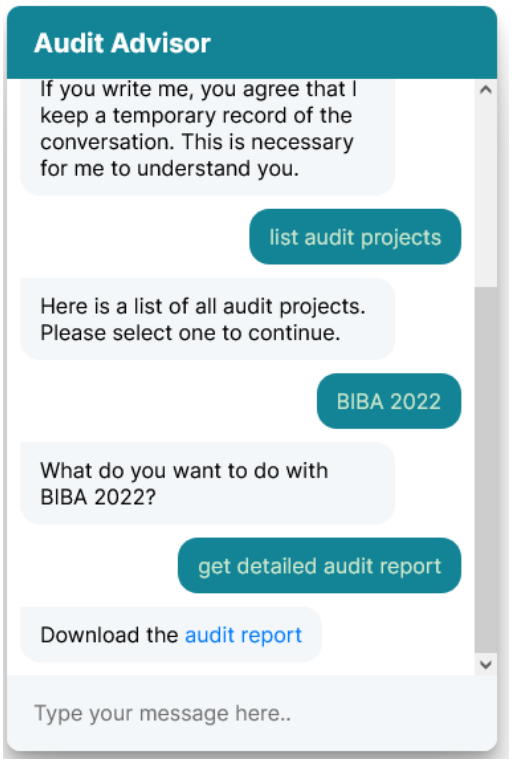

Figure 25. Bot provides audit preparation report

The basic implementation focuses on user input as free text or URLs. This input type is readily available in Rasa. Advanced implementations could allow file uploads to collect evidence (e.g., a PDF document with an ISO 9001:2015 certificate). These features require customization of the chat widget, a document storage service, and additional functions in the action server. Outputs can be text, links, buttons, and images.

6. Conclusions

The goal of i4Q^{LCP} is to ensure data quality in smart manufacturing by providing recommendations for certification and audit procedures. This goal has been achieved using relevant standards in quality management and data quality. The i4Q^{LCP} provides a procedure to establish, maintain, and audit a high data quality standard needed for the use of innovative technologies like the i4Q solutions.

D5.12 presents the second version (v2) of i4Q^{LCP} Solution. In this second version, the audit guideline has been finalized and the digital assistant has been created. The document includes parts of D5.6, the first version of this guideline solution. There are descriptions of audit scope, objectives, risks and opportunities, audit team roles, resource planning, audit methods, documentation, implementation, conduct, and improvement included that have been defined according to ISO 19011:2018 [6]. Furthermore, the development of audit criteria has been completed. This enables to evaluate data quality according to ISO/IEC 25012:2008 [8] including the description of DLC and ISO/IEC 25024:2015 [9] with the related data quality characteristics. Other standards have been considered to define important prerequisites for i4Q^{LCP} in manufacturing environment, namely ISO 10012:2003 [2], ISO 15939:2017 [5] as well as ISO 8000 [11-19]standards series and ISO/IEC/IEEE 24765:2017 [7] providing important data quality basic knowledge and vocabulary. i4Q^{LCP} guideline should enable users to create an audit plan for auditing data quality of a specific manufacturing device and follow the audit criteria by using the IT-based assistant tool.

To evaluate if the audit criteria have been met, questions related to each audit criteria have been created. These questions are integrated in the IT-based assistant tool. This tool consists of a chatbot which guides auditees and auditors in determining if audit criteria have been met. These audit criteria questions are developed based on target entities and properties of data as defined in ISO/IEC 25024:2015 [9] and follow the data life-cycle phases according to ISO/IEC 25012:2008 [8]. The tool also provides an automatic evaluation of the level of conformity.

The validation of the i4Q^{LCP} IT-based assistant tool and the guideline document will be done with i4Q partners and pilots within the WP6 activities. The i4Q^{LCP} audit procedure provides a basis for data quality certification processes and standardisation activities to complement existing quality certifications. This is elaborated in WP7.

Since Data Quality is an essential aspect of data collection, transformation, and analytics within the industrial manufacturing environment, this guideline is an important part of the i4Q framework. By disclosing deficiencies in the assurance of data quality, recommendations for reconfiguration of process, audit strategies, certification, and regulation have been offered. This can support to enhance the outcomes of i4Q tasks T5.1 to T5.4.

In summary it is stated that the three requirements for i4Q^{LCP} defined in WP1 could be fully met. Namely they request 1) an audit procedure for manufacturing resources (machine, cell or manufacturing line) to ensure the accuracy and reliability of data from the manufacturing process, 2) knowledge transfer to the employees through training according to guidelines and applied standards, and 3) an IT-assistant to support audit teams to conduct the audit procedure.

7. References

[1] Glenn.N.; Understand Data Analysis, Analyst Answers, https://analystanswers.com/what-is-a-data-attribute-definition-types-examples/#:~:text=In%20short%2C%20a%20data%20attribute,programming%20languages%20such%20as%20Python (Last accessed 4 May 2022)

[2] ISO 10012 Measurement Management Systems – Requirements for measurement processes and measuring equipment (ISO 10012:2003)

[3] ISO/IEC/IEEE 12207 Systems and Software Engineering – Software life cycle processes (ISO/IEC/IEEE 12207:2017)

[4] ISO/IEC/ IEEE 15288 Systems and Software Engineering – System life cycle processes (ISO/IEC/IEEE 15288:2015)

[5] ISO/IEC/IEEE 15939 Systems and software engineering – Measurement process (ISO /IEC/IEEE 15939-2017(E))

[6] ISO 19011 Guidelines for auditing management systems (ISO 19011:2018)

[7] ISO/IEC/IEEE 24765 Systems and Software Engineering- Vocabulary (ISO/IEC/IEEE 24765:2017(E))

[8] ISO/IEC 25012 Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model (ISO/IEC 25012:2008(E))

[9] ISO/IEC 25024 Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) - Measurement of data quality (ISO/IEC 25024:2015(E))

[10] ISO IEC 33020 Information technology - Process assessment - Process measurement framework for assessment of process capability (ISO IEC 33020:2019)

[11] ISO 8000-02 Data quality – Part 2: Vocabulary (ISO 8000-2:2020(E))

[12] ISO 8000-08 Data quality Part 8: Information and data quality: Concepts and measuring (ISO 8000-8: 2015(E))

[13] ISO 8000-61 Data quality Part 61: Data quality management: Process reference model (ISO 8000-61:2016(E))

[14] ISO 8000-62 Data quality Part 62: Data quality management: Organizational process maturity assessment: Application of standards relating to process assessment (ISO 8000-62:2018(E))

[15] ISO 8000-63 Data quality Part 63: Data quality management: Process measurement (ISO 8000-63:2019(E))

[16] ISO/DIS 8000-64 Data quality Part 64: Data quality management: Organizational process maturity assessment: Application of the Test Process Improvement method (ISO/DIS 8000-64:2021(E))

[17] ISO/TS 8000-65 Data quality Part 65: Data quality management: Process measurement questionnaire (ISO/TS 8000-65:2020(E))

[18] ISO/TS 8000-81 Data quality Part – 81 Data quality assessment: Profiling (ISO/TS 8000-81:2021(E))

[19] ISO/TS 8000-150 Data quality – Part 150: Master Data: Quality management framework (ISO/TS 8000-150:2011(E))

[20] ISO 9000 Quality Management Systems – Fundamentals and Vocabulary (DIN EN ISO 9000:2015 –11)

[21] ISO 9001 Quality Management systems- Requirements (ISO 9001:2015)

[22] ISO Survey 2020 https://www.iso.org/the-iso-survey.html (Last accessed 30 May 2022)

[23] Joachim Herrmann and Holger Fritz, 2021: Qualitätsmanagement - Lehrbuch für Studium und Praxis, 3., aktualisierte und erweiterte Auflage, Hanser Munich

[24] Nowak-Meitinger, A. M.; Mayer, J. P.; Jochem, R. (2021). i4Q – *Deliverable 1.9 - Requirements Analysis and Functional Specification v2*. (Submitted 31-09-2021)

[25] Rahul, K., & Banyal, R. K. (2020). Data life cycle management in big data analytics. *Procedia Computer Science*, *173*, 364-371.

[26] Rasa, https://rasa.com/docs/rasa/connectors/your-own-website/ (Last accessed 25 May 2022)

[27] SEVOCAB (2022). Software and Systems Engineering Vocabulary. Link: https://pascal.computer.org/sev_display/search.action (Last accessed 25 May 2022)

[28] Singh, A. (2022). Data Documentation, i2i dimewiki world bank. Link: https://dimewiki.worldbank.org/Data_Documentation#:~:text=Data%20documentation%20is%20the%20process,course%20of%20a%20research%20project (Last accessed 9 November 2022)

[29] Stobierski, T. (2021); 8 Steps in the data life cycle, Harvard Business School Online. Link: https://online.hbs.edu/blog/post/data-life-cycle (Last accessed 9 November 2022)

[30] Techopedia, https://www.techopedia.com/definition/8/database-column (Last accessed 4 May 2022)

[31] TechTarget, https://www.techtarget.com/whatis/definition/format#:~:text=Also%20see%20file%20format.,the%20same%20or%20another%20format (Last accessed 4 May 2022)

[32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

8. Appendix

i4Q Manufacturing Line Data Certification Procedure (i4Q^{LCP}) web documentation can be accessed online at: https://i4q.upv.es/22_i4Q_LCP .

8.1 Appendix 1: Requirements and Functional Specification of i4Q^{LCP}

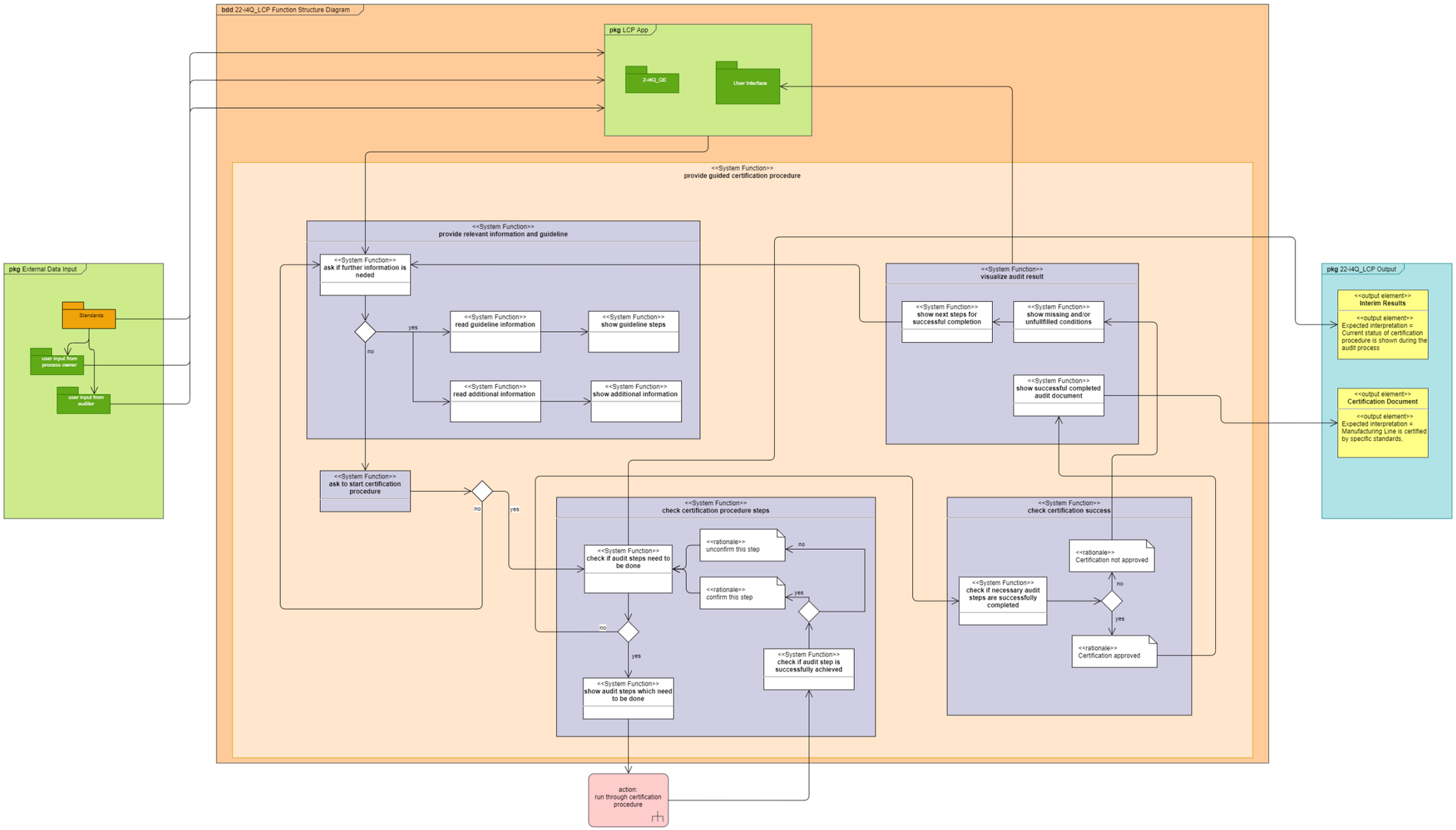

Figure 26. i4Q^{LCP} Function Structure Diagram (FSD) (D1.9)[21]

[21] This figure is taken from i4Q Deliverable D1.9 [24] (Nowak-Meitinger et al., 2021, p. 121)

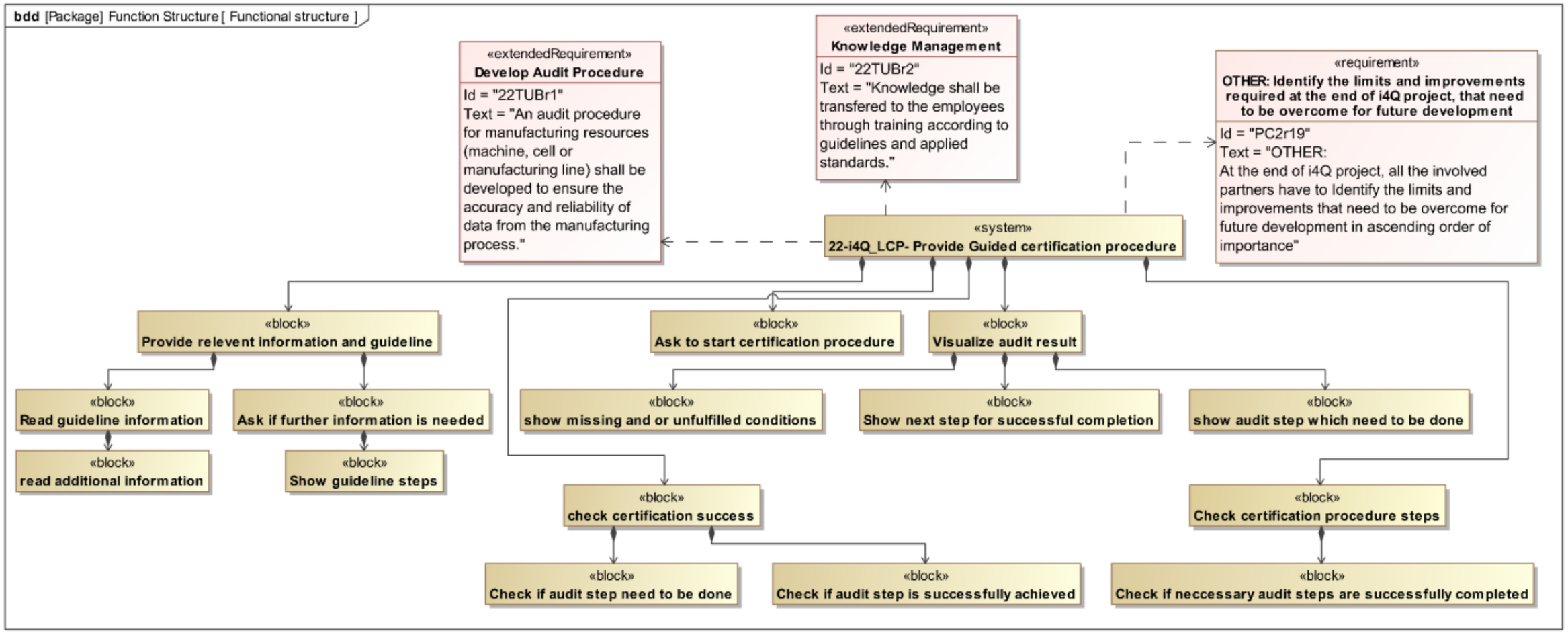

Figure 27. i4Q[LCP] Requirements Mapping and Functional Specification (D1.9)[22]

[22] [24] This figure is taken from i4Q Deliverable D1.9 (Nowak-Meitinger et al., 2021, p. 122)

8.2 Appendix 2: Terms and definitions related to audits

The terms and definitions in **Table 12** are based on ISO 19011:2018. These will be used during the development of i4Q[LCP].

Term	Definitions (ISO 19011:2018, ISO 9000:2015-11)
Audit	Systematic, independent, and documented process for obtaining objective evidence and evaluating it objectively to determine the extent to which the audit criteria are fulfilled.
Audit client	Organization or person requesting an audit; In the case of internal audit, the audit client can also be the auditee or the individual(s) managing the audit programme.
Audit conclusion	Outcome of an audit, after consideration of the audit objectives and all audit findings.
Audit criteria	Set of requirements used as a reference against which objective evidence is compared.
Audit evidence	Records, statements of fact or other information, which are relevant to the audit criteria and verifiable.
Audit findings	Results of the evaluation of the collected audit evidence against audit criteria.
Audit plan	Description of the activities and arrangements for an audit.
Audit programme	Arrangements for a set of one or more audits planned for a specific time frame and directed towards a specific purpose.
Audit scope	Extent and boundaries of an audit.
Audit team	People conducting an audit, supported if needed by technical experts; One auditor of the audit team is appointed as the audit team leader.
Auditee	Organization as a whole or parts thereof being audited.
Auditor	Person who conducts an audit.
Certification	Third-party attestation related to an object of conformity assessment, with the exception of accreditation (DIN EN ISO/IEC 17000:2020-09)
Competence	Ability to apply knowledge and skills to achieve intended results.
Conformity	Fulfilment of a requirement.
Effectiveness	Extent to which planned activities are realized and planned results achieved.
Management system	Set of interrelated or interacting elements of an organization to establish policies and objectives, and processes to achieve those objectives.
Nonconformity	Non-fulfilment of a requirement.
Objective evidence	Data supporting the existence or verity of something; Objective evidence can be obtained through observation, measurement, test or by other means; Objective evidence for the purpose of the audit generally consists of records, statements of fact, or other information which are relevant to the audit criteria and verifiable.
Observer	Individual who accompanies the audit team but does not act as an auditor.

Performance	Measurable result.
Process	Set of interrelated or interacting activities that use inputs to deliver an intended result.
Requirement	Need or expectation that is stated, generally implied or obligatory.
Risk	Effect of uncertainty.
Technical expert	<audit> person who provides specific knowledge or expertise to the audit team. Notes: Specific knowledge or expertise relates to the organization, the activity, process, product, service, discipline to be audited, language or culture. A technical expert to the audit team does not act as an auditor.

Table 12. Terms and Definitions for i4Q[LCP] related to audits[23]

[23] [32] Trevino, R.; Nowak-Meitinger, A. M.; Wellsandt, S. (2022). i4Q – *Deliverable 5.6 - Manufacturing Line Data Certification Procedure*. (Submitted 30-06-2022)

8.3 Appendix 3: Terms and Definitions related to data quality

To ensure interoperability with existing standards, the vocabulary used in the audit procedure of i4Q[LCP] is based on the terms and definitions of *ISO 8000-02:2020 - Data quality – Part 2: Vocabulary* and *ISO/IEC/IEEE 24765:2017 Systems and software engineering — Vocabulary* (**Table 13**). ISO/IEC/IEEE 24765:2017 introduces a vocabulary and terminology which is consistent for all ISO standards within the scope of *Software and systems engineering* and the IEEE Computer Society and Software Engineering Standards Committee.

Term	Definitions
Accessibility	Extent to which products, systems, services, environments and facilities can be used by people from a population with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use [7] The degree to which data can be accessed in a specific context of use, particularly by people who need supporting technology or special configuration because of some disability [8]
Accuracy	The degree to which data has attributes that correctly represent the true value of the intended attributes of a concept or event in a specific context of use [8]
Availability	The degree to which data has attributes that enable it to be retrieved by authorized users and/or applications in a specific context of use [8]
Completeness	The degree to which subject data associated with an entity has values for all expected attributes and related entity instances in a specific context of use [8]
Compliance	Doing what has been asked or ordered, as required by rule or law [7]. The degree to which data has attributes that adhere to standards, conventions or regulations in force and similar rules relating to data quality in a specific context of use [8]
Confidentiality	The degree to which data has attributes that ensure that it is only accessible and interpretable by authorized users in a specific context of use [8]
Consistency	The degree to which data has attributes that are free from contradiction and are coherent with other data in a specific context of use. It can be either or both among data regarding one entity and across similar data for comparable entities [8]
Column	A set of related data items in a table24
Credibility	The degree to which data has attributes that are regarded as true and believable by users in a specific context of use [8]
Currentness	The degree to which data has attributes that are of the right age in a specific context of use [8]

[24] https://www.techopedia.com/definition/8/database-column

Term	Definitions
Data	Reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing [11]
Data Accuracy	Quality of data in respect of the represented value agreeing with the corresponding true value to a degree necessary for an intended purpose [11]
Data Attribute	Smallest parcel of information, within an identified data group, carrying a meaning from the perspective of the software's functional user requirements/header of column in table [7] and [25]
Data File	A set of related data records treated as a unit [7]
Data Format	An arrangement of data for storage and display[26]
Data Items	The smallest identifiable unit of data within a certain context for which the definition, identification, permissible values, and other information are specified by means of a set of properties. Common five types are Integer, Date, Time, Text, and Boolean [7]
Data Processing	Systematic performance of operations upon data [7]
Data Quality	Degree to which the characteristics of data satisfy stated and implied needs when used under specified conditions [7]
Data Record	A set of related data items treated as a unit [7]
Data Set	Logically meaningful grouping of data (ISO 8000-02:2020)
Data Store	Organized and persistent collection of data and information that allows for its retrieval [7]
Data Type	Categorization of an abstract set of possible values, characteristics and set of operations for an attribute [7]
Data Value	Content of data item [7]
Database management system	Organized collection of structured data (ISO/IEC 25024:2015 Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Measurement of data [7]
Defined process	Implemented process that is managed and tailored from the organization's set of standard processes according to the organization's tailoring guidelines [7]
Description	Information item that represents a planned or actual concept, function, design, or object [7]
Documentation	Collection of documents on a given subject, written or pictorial information describing, defining, specifying and reporting [7]
Domain	Distinct scope, within which common characteristics are exhibited, common rules observed, and over which a distribution transparency is preserved [7]

[25] https://analystanswers.com/what-is-a-data-attribute-definition-types-examples/#:~:text=In%20short%2C%20a%20data%20attribute,programming%20languages%20such%20as%20Python

[26] [31] https://www.techtarget.com/whatis/definition/format#:~:text=Also%20see%20file%20format.,the%20same%20or%20another%20format

Term	Definitions
Efficiency	The degree to which data has attributes that can be processed and provide the expected levels of performance by using the appropriate amounts and types of resources in a specific context of use [8]
Maturity Model	Capability level at which the organization is performing the processes for data quality management [11]
Master Data	Data held by an organization to describe the entities that are both independent and fundamental for that organization, and referenced in order to perform its transactions [11]
Metadata	Data defining and describing other data [11]
Measure	Ascertain or determine the magnitude or quantity of something [11]
Measurement	Result of measuring something [11]
Portability	The degree to which data has attributes that enable it to be installed, replaced or moved from one system to another preserving the existing quality in a specific context of use [8]
Precision	The degree to which data has attributes that are exact or that provide discrimination in a specific context of use [8]
Product Data Quality	Thing or substance produced by a natural or artificial process degree to which a set of inherent characteristics of data fulfils requirements [11]
Recoverability	Degree to which, in the event of an interruption or a failure, a product or system can recover the data directly affected and re-establish the desired state of the system [7] The degree to which data has attributes that enable it to maintain and preserve a specified level of operations and quality, even in the event of failure, in a specific context of use [8]
Technical expert	Person who provides specific knowledge or expertise to the audit team [11]
Traceability	The degree to which data has attributes that provide an audit trail of access to the data and of any changes made to the data in a specific context of use [8]
True Value	Value that characterizes a characteristic perfectly defined in the conditions that exist when the characteristic is considered [7]
Understandability	Ease with which a system can be comprehended at both the system-organizational and detailed-statement levels [7] The degree to which data has attributes that enable it to be read and interpreted by users, and are expressed in appropriate languages, symbols, and units in a specific context of use [8].

Table 13. Terms and Definitions for i4Q[LCP] related to data quality

8.4 Appendix 4: Audit Criteria Questions

Target Entity (ISO/IEC 25024:2015(E))	Properties (ISO/IEC 25024:2015(E))	Description (ISO/IEC 25024:2015(E))	Data Quality Characteristics (ISO/IEC/IEEE 25012:2008).	ID	Description (ISO/IEC/IEEE 25012:2008).	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task
DBMS Database Management System	Data Format	arrangement of data (4.5) for storage or display	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Are the standards, conventions or regulations for external data values compliant to the internal data collection standards and regulations?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes	
							No	Data values shall be defined and documented according to external/internal standards, conventions, and regulations.		No	Standards, conventions or regulations for external data need to be compliant to standards and regulations for the internal data use.		No	Define, establish and document a process for use of internal/external standards, conventions, and regulations for data arrangement.		No	Provide related persons access to documentation of all used internal/external standards, conventions or regulations.
DBMS Database Management System	Data Item	smallest identifiable unit of data (4.5) within a certain context for which the definition, identification, permissible values, and other information (4.21) is specified by means of a set of properties	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Are the standards, conventions or regulations for external data values compliant to the internal data collection standards and regulations?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes	
							No	Data values shall be defined and documented according to external/internal standards, conventions, and regulations.		No	Standards, conventions or regulations for external data need to be compliant to standards and regulations for the internal data use.		No	Define, establish and document a process for use of internal/external standards, conventions, and regulations for data arrangement.		No	Provide related persons access to documentation of all used internal/external standards, conventions or regulations.
DBMS Database Management System	Data Item	smallest identifiable unit of data (4.5) within a certain context for which the definition, identification, permissible values, and other information (4.21) is specified by means of a set of properties	Confidentiality	Cnf-D-1	Degree to which data item defined as confidential can be accessed by authorized users only	Does your company store confidential data that only authorized users should be able to access?	Yes		Does your company use processes ensuring confidential data is only accessible by authorized personnel?	Yes		Is the documented process available to relevant personnel?	Yes				
							No	No action required		No	Processes to ensure confidential data is only accessible by authorized personnel shall be defined, documented and implemented.		No	Access shall be provided to the relevant personnel.			

Table 14. Examples of Audit Criteria Questions for DLC- Data Integration

Target Entity (ISO/IEC 25024:2015(E))	Properties (ISO/IEC 25024:2015(E))	Description (ISO/IEC 25024:2015(E))	Data Quality Characteristics (ISO/IEC/IEEE 25012:2008)	ID	Description (ISO/IEC/IEEE 25012:2008).	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task
Form	Data Format	arrangement of data (4.5) for storage or display	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Are the standards, conventions or regulations for external data values compliant to the internal data collection standards and regulations?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes	
							No	Data values shall be defined and documented according to external/internal standards, conventions, and regulations.		No	Standards, conventions or regulations for external data need to be compliant to standards and regulations for the internal data use.		No	Define, establish and document a process for use of internal/external standards, conventions, and regulations for data arrangement.		No	Provide related persons access to documentation of all used internal/external standards, conventions or regulations.
Form	Data Format	arrangement of data (4.5) for storage or display	Accessability	Acs-D-2	Degree to which data or information are not accessible by the intended users due to a specific format	Does your company ensure accessibility of all formats of data or information used in the company to the intended users (e.g. through software, systems, etc.)?	Yes		Does your company use processes to ensure accessibility of data or information for intended users?	Yes		Is the documented process available to relevant personnel?	Yes				
							No	Accessibility of all formats of data or information used in the company to the intended users (e.g. through software, systems, etc.) shall be ensured.		No	Processes ensuring accessibility shall be defined, implemented and documented.		No	Access shall be provided to the relevant personnel.			
Form	Data Item	smallest identifiable unit of data (4.5) within a certain context for which the definition, identification, permissible values, and other information (4.21) is specified by means of a set of properties	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Are the standards, conventions or regulations for external data values compliant to the internal data collection standards and regulations?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes	
							No	Data values shall be defined and documented according to external/internal standards, conventions, and regulations.		No	Standards, conventions or regulations for external data need to be compliant to standards and regulations for the internal data use.		No	Define, establish and document a process for use of internal/external standards, conventions, and regulations for data arrangement.		No	Provide related persons access to documentation of all used internal/external standards, conventions or regulations.

Table 15. Examples of Audit Criteria Questions for DLC- Presentation

Target Entity (ISO/IEC 25024:2015(E))	Properties (ISO/IEC 25024:2015(E))	Description (ISO/IEC 25024:2015(E))	Data Quality Characteristics (ISO/IEC/IEEE 25012:2008)	ID	Description (ISO/IEC/IEEE 25012:2008)	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task	Questions for Guideline	Ans	Task
RDBMS Relational Data Base Management System	Data Format	arrangement of data (4.5) for storage or display	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Are the standards, conventions or regulations for external data values compliant to the internal data collection standards and regulations?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes	
							No	Master data should be stored in a clear and systematic manner.		No	Metadata describing the master data shall be stored in a systematic and clear manner so that the master data is easily understandable.		No	Processes to ensure master data's´ understandability shall be defined, implemented, and documented.		No	Access shall be provided to the relevant personnel.
RDBMS Relational Data Base Management System	Data Item	smallest identifiable unit of data (4.5) within a certain context for which the definition, identification, permissible values, and other information (4.21) is specified by means of a set of properties	Consistency	Con-I-1	For each value of one attribute of a table exists the same value of the same attribute in a different table; i.e. there is link between the same attribute represented in different tables and they contain the same values	Is the data from one file also stored/used in another file?	Yes		Does your company use processes ensuring that the data in both files are identical and consistent?	Yes		Is the documented process available to relevant personnel?	Yes				
							No	No action required.		No	Processes shall be defined, implemented and documented that ensure the data in both files are identical and consistent.		No	Access shall be provided to the relevant personnel.			
RDBMS Relational Data Base Management System	Data Item	smallest identifiable unit of data (4.5) within a certain context for which the definition, identification, permissible values, and other information (4.21) is specified by means of a set of properties	Compliance	Cmp-I-1	Degree to which data values and/or format comply with specific standards, conventions or regulations	Does your company use internal/external standards, conventions or regulations to arrange external data for storage or display ?	Yes		Are the standards, conventions or regulations for external data values compliant to the internal data collection standards and regulations?	Yes		Is the use of internal/external standards, conventions, and regulations documented and communicated to all relevant persons?	Yes		Do relevant persons have access to all used internal/external standards, conventions or regulations?	Yes	
							No	Data shall be encrypted as per internal/external standards, customer requirements, conventions or regulations.		No	Process to ensure that data is encrypted shall be documented and implemented.		No	Process documentation to be made available for personnel.		No	Process documentation to be made available for personnel.

Table 16. Examples of Audit Criteria Questions for DLC- Data Store