TURNING MAINTENANCE REPORTS INTO ACTIONABLE RISKS AND WORK ORDERS WITH LLM

Paper No. EUR26_04 PCIC energy

Tomas Kozel
ABB s.r.o.
Videnska 117, Brno
Czech Republic

Mariam Ibrahim
ABB B.V.
G. Hintzenweg 81, Rotterdam
Netherlands

Giada Volpin
ABB S.p.A.
Via Friuli 4, Dalmine
Italy

Alberto Carini
ABB S.p.A.
Via Friuli 4, Dalmine
Italy

Abstract - Three years ago, we conducted market research to understand gaps maintainers see in the maintenance of electrical distribution equipment. The loss of information captured in maintenance reports was identified as key pain point and very common hurdle. After a job is completed, a report is stored - physically or digitally - as evidence that the device was inspected or repaired. These reports often contain important observations and recommendations for follow-up actions, yet they are rarely reviewed and therefore these practical insights from field go unused.

To address the gap, we leveraged a Large Language Model (LLM) to identify key observations and recommendations in the reports. It automatically analyzes maintenance reports and presents, in a structured way, the observations, risks and recommendations distilled from them. The system addresses the common pain point for maintainers and complements existing systems by ensuring that field recommendations are collected and reported risks can be evaluated and acted upon.

Index Terms — Maintenance reports processing, Data mining, Large Language Models, Asset risk management

I. INTRODUCTION

In 2023, market research was conducted to identify gaps perceived by maintenance managers and operators in the maintenance process of their electrification equipment. The research revealed that equipment owners typically have well-established processes for identifying general maintenance plans, dispatching technicians or service providers, and closing activities with reports.

A. Gaps Identified

Although the overall process showed no major gaps, two common issues in the maintenance process were identified:

1) *Missing Prioritization of Maintenance Actions*: While maintenance actions might be clearly identified, a prioritization process is sometimes missing. To achieve ever-increasing targets in operational availability and reliability while reducing the maintenance budget, maintenance managers need to identify the most impactful actions with the highest risk reduction and avoid costly maintenance with low impact on system performance.

2) *Information from Reports Is Lost in Repositories:* Reports are typically stored in physical or digital repositories without any follow-up on the details contained within. Many reports include important observations and recommendations for further maintenance actions, which are thus lost and not acted upon, leading to further maintenance suboptimization.

We targeted these two gaps in our further work, conceiving concepts and technology application proposals to address them.

B. Observation, Risk, Recommendation, Action

To support our further work, an overarching concept of Observation, Risk, Recommendation, Action (ORRA) was created to describe the ideal maintenance process – see Fig.1.

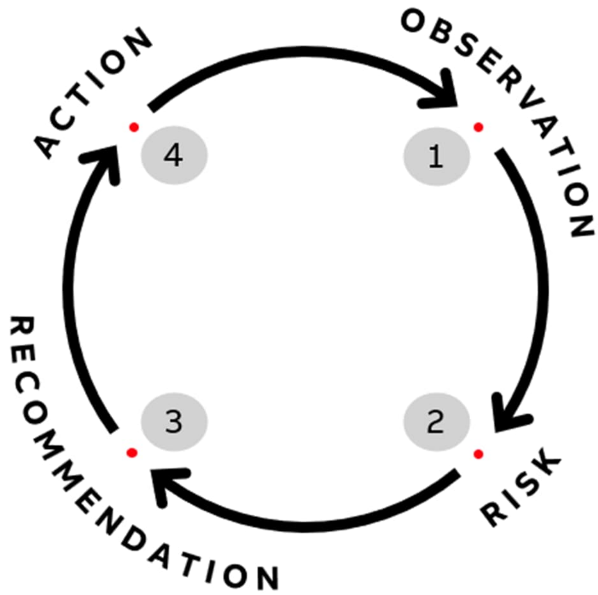


Fig. 1 Ideal maintenance process visualization

1) *Ideal Process Description*: The process ideally starts with initial observations (1) created during a periodic inspection or similar planned activity. These observations serve as crucial inputs for calculating the reliability risk (2) for a device. Once the risk is defined, a recommendation (3) can be formulated as a proposal to mitigate the risk. At this point, the maintenance manager should have a list of quantified risks for the devices under their responsibility, select the top risks, and plan associated recommendations for fulfilment. The recommendations become actions (4) once they are planned and field service personnel are assigned to execute them. Once the actions are executed, the report of the activity is taken as a new input for another loop. Observations from the report are used to update the risk calculations, and subsequently, new risks and recommendations might appear in the maintenance

manager's list and get assigned for execution as actions.

2) *Current Gaps in the Process:* the gaps identified in the research are currently disrupting this process in two ways, as depicted on Fig.2

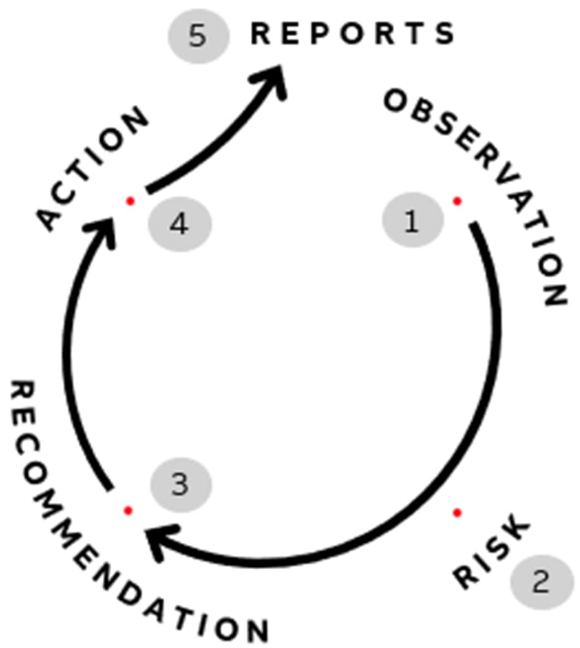


Fig. 2 Current maintenance process with gaps disrupting the efficiency

The first visible gap is skipping the risk quantification (2) and deriving recommendations (3) directly from observations. In this case, the recommendations cannot be presented in a prioritized manner, from those mitigating the highest risks to less important ones. The action planning is then also affected, and actions (4) with low effect on system reliability might be attended to first, keeping the field service crew busy inefficiently.

The second gap is depicted as an arrow leading from Action (4) out of the circle to Reports (5) instead of closing the circle. If stored in the repository without proper extraction of observations, the reports of field service activity do not bring any new observations into the loop. The process stops and doesn't receive any new observations to process until a dedicated inspection activity in the future discovers that there is still something to attend to, even though it was previously reported.

II. ADDRESSING THE GAPS IDENTIFIED BY THE RESEARCH

A. *The Risk Quantification Gap Mitigation*

Further exploration of the gap of skipping risk quantification revealed that only small operators with limited number of electrical distribution devices and without well-organized maintenance programs and crews face this issue.

Operators with large installed base or with critical processes to keep in operation typically use an Enterprise Asset Management System (EAMS) for registering, evaluating and managing the risks associated with power distribution infrastructure in their facilities. The smaller operators, however, do not consider standard EAMS systems a good fit for their needs due to the complexity of system implementation and high operational costs.

This situation opens space for small-scale Asset Management Systems (AMS), that do not cover a wide variety of assets with all the associated complexity but rather concentrate on power distribution systems and components. Equipment manufacturers of the devices prevailing in the operator's electrical distribution system might be well positioned for offering such small-scale AMS, provided they have an installed base of their products on customer sites mapped in a system.

The installed base map might serve as source of truth for the small-scale AMS, avoiding the time-consuming registration and attribution needed for dedicated EAMS. The small-scale AMS might be then created as a customer portal on top of the data, allowing observation, risk, recommendation and action handling. Utilizing historical maintenance reports as data source, no additional hardware (such as gateways or sensors) is required, and the solution is applicable to all types of assets, including non-digital ones. Easy access to device documentation and maintenance history might complement the system features to close the loop by providing 360° view on the electrical distribution assets.

In summary, the problem of skipping risk quantification for small operators is solvable by implementing EAMS or small-scale AMS systems available on the market and thus does not require further research or technology exploration.

B. *The Insights Lost in Reports Gap Mitigation*

The situation is different for the second gap identified in the research - the valuable observations from the field being lost in report repositories without any follow-up actions. This gap was identified as significant efficiency disturbance by all the operators interviewed, independent of their size or whether they operate in the industrial or residential segment.

Operators with a large installed base might have a vast number of reports to be reviewed and scanned for potential findings and recommendations. As the process of reviewing reports is manual and the number of pages might be significant, it would be easy to overlook important issues mentioned in the report. The problem is prominent with reports documenting periodic tests when the reports do not include a high-level summary section where all the important observations and recommendations would be summarized.

For smaller systems operators, power distribution system maintenance is typically under-resourced or outsourced. In such a setup, we frequently see every player considering the report review and extraction of information as duty of the other party, resulting in no review at all.

In many cases, the formal report review system is replaced by informal face-to-face conversations between field operators and maintenance managers. The main hurdles encountered on the job are typically shared, which might include valuable information on potential findings and recommendations. Many of them, however, are lost in the word-of-mouth nature of this informal channel.

The exploration of automation in the field of data mining from reports didn't reveal any existing tools or systems capable of solving the situation. The technology exploration, nonetheless, showed a possible direction for automation.

In 2023 the generative AI in the form of Large Language Models (LLMs) experienced an unprecedented boom. The promise of the technology included natural language text processing, document summarization, and important information extraction. Therefore, it was only natural to start exploring the use of LLMs in the context of automating report data mining.

C. *Utilizing LLMs for Text Mining: A Grounded, Modular, and Human-based Validation Framework*

Even though LLMs offer a promising path to mine free text in reports and automatically extract structured knowledge, the straightforward application of such technology in the industrial domain faces several challenges. General-purpose LLMs typically struggle with handling specialized terminology and context, making them impractical as off-the-shelf models for domain specific information extraction. In addition, unconstrained LLM outputs can produce factual errors (called hallucinations), which can adversely impact trust and reliability - both crucial in industrial contexts. To overcome these challenges, we structured our approach for leveraging LLMs around 3 core principals

1) *Grounding prior to generation:* LLMs must be grounded by a domain data model that provides them with information regarding what to look for, what to ignore, and how to provide the results.

2) *Platform over model:* The solution needs to be model-agnostic and adaptable to the rapid changes in the AI landscape as well as the diverse nature of the reports that we anticipate receiving.

3) *Human-in-the loop by design:* Expert validation is essential to achieve industrial-grade quality and trust.

The diagram below provides a high-level overview of the approach, with key components detailed in the following sections.

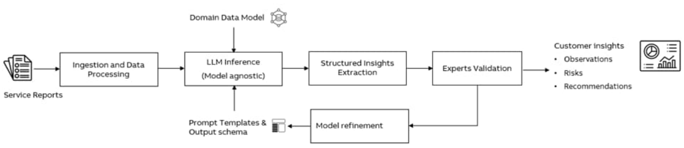


Fig. 3 Overview of LLM-based text mining framework

D. *Grounding LLMs with a Domain Data Model*

Instead of asking an LLM to "interpret" service reports freely, the framework grounds the model using a Domain Data Model that defines:

- The relevant entities (e.g., equipment, observation, risk, recommendation, action, etc.)
- The permitted relations (e.g., mitigated by, caused by, etc.)
- The controlled vocabulary and domain concepts

This grounding helps reduce hallucinations, improves consistency across reports, and supports explainable, auditable results. The LLM, therefore, becomes a trusted information extraction engine, as opposed to a text generator that produces creative content.

E. *Flexible Template-Based Mining Platform*

To be sustainable and future-proofed, the framework separates the functions of what the model does from the models that can be used to perform those functions. The following are the main characteristics for enabling this separation:

- Prompt templates that are aligned with specific extraction tasks (e.g., "extract observed issues," "identify failure modes mentioned," etc.)
- Structured output schemas that enable tying the output to the data model definitions
- Pluggable LLM interface to allow for easy replacement or upgrade of the models
- Batch and incremental processing capabilities to facilitate historical backfills and ongoing ingestion

Such a design paves the way for fast experimentation with multiple models to stay aligned with the rapid advancement in technology without the need to re-engineer the platform.

F. *Human-in-the-Loop Validation and Continuous Improvement*

Industry-grade accuracy cannot be achieved through AI-based automation alone. Therefore, our framework embeds a consistent expert validation process for the extracted data for the following purposes:

- Flagging low-confidence extractions
- Allowing experts to review, correct or reject the extracted output
- Capturing the experts' feedback as structured signals for model refinement

This approach supports continuous improvement of extraction, as well as trust and transparency for user-facing risks and recommendations.

G. *Summary of Initial Experimentation Results*

Our goal was to design a robust prompt that would effectively extract information from a large and diverse corpus of industrial maintenance documentation and service reports. For this purpose, a set of 80 service reports (multilingual, narrative, and table-based) was considered in the prompt design phase to ensure that it would generalize across the wide variety of real-world reporting situations. The testing process utilized a set of 10 reports and was manually evaluated in terms of overall system performance as well as how the system performed with respect to different reporting formats (e.g., narrative, table). For LLM inference, Microsoft Azure OpenAI models were leveraged, along with document handling using Azure Document Intelligence, which showed promising results across a broad range of file formats.

Overall, the extraction accuracy was found to be approximately 90%, indicating strong initial performance. An example of an input report and the results of the data mining are shown in the Appendices A and B.

The testing also indicated that the structure of the report significantly impacted the ability to extract data properly; notably, reports that contained tables were found to be much more difficult to parse correctly. In response to these challenges, several new specialized parsing modules for nested tables and hybrid layouts were developed.

Additionally, the testing of multilingual reports indicated that the varied performance of the LLM is based on the clarity and structure of the language used in the report. Service reports containing ambiguous language resulted in decreased levels of extraction accuracy. Accordingly, targeted efforts were deployed to standardize report-writing templates and train service professionals to provide consistent and clear semantics in future service documentation, which led to significant improvements in the extraction results from the models.

Further improvements in extraction performance were seen through the structured gathering of feedback from subject matter experts on the quality of extraction. Their evaluation highlighted the model limitations in extraction quality and were used for informed prompt refinement. This feedback loop will serve as a basis for continuous model refinement towards industrial-grade reliability and accuracy at scale.

III. CONCLUSIONS

This paper describes the gaps in maintenance process of electrical distribution equipment identified by a market research in 2023. It suggests addressing the gap of insights from field maintenance reports being lost in repositories by forming a data mining system based on Large Language Model technology. The three main challenges and associated design principles used in the development of a prototype solution are described and explained in the paper. The idea of utilizing Large Language Models for the task is supported by a set of initial experiments that confirm the value of this approach.

There are several aspects influencing the capability of the prototype to extract valuable information. Reports containing tables are harder to parse for high-quality extraction. The clarity and non-ambiguity of the language used in the maintenance reports is identified as another strong influencing factor.

The prototype system is capable of mining the report information with success rate above 90% for the reports that are not affected by the above-mentioned quality-limiting aspects. This result gives good confidence for the pilot and gradually wide-scale deployment.

IV. ACKNOWLEDGEMENTS

Special credits and thanks to Andrzej Goszczynski, ABB Sp. z o.o. for all the hard work on the report mining pipeline. Your work gave us confirmation this approach is valuable.

V. VITA

Tomas Kozel graduated from the Technical University in Brno in 2001 with an Engineering degree in Power Systems. He has taken different roles in ABB in the field of MV switchgear servicing since 2003 and has contributed to applying AR and AI technologies to electrification equipment servicing use cases since 2017.
tomas.kozel@cz.abb.com

Mariam Ibrahim graduated with an MPhil in Energy Technologies from the University of Cambridge in 2017, and with bachelor's degree in Mechatronics Engineering with Highest Honors from the German University in Cairo in 2016. She currently serves as Global Product Owner of Data and AI in ABB Electrification Service Division. Her work focuses on digital and data-driven solutions for energy systems.
Mariam.ibrahim@nl.abb.com

Giada Volpin graduated from Ca' Foscari University of Venice with a master's degree in international management and conducted award-winning research in user behavior at Capital Normal University of Beijing. She is a Global Product Manager at ABB Electrification Service, where she leads the digitalization of technical support, repair, and training services.
Giada.volpin@it.abb.com

Alberto Carini graduated from ITIS Paleocapa of Bergamo with an Electrical Engineering high school diploma.
He has held different roles in Marketing & Sales and product management for MV and LV switchgear. He is currently Global Product Manager for brownfield digitalization
alberto.carini@it.abb.com

APPENDIX A

Sample Service Report

Service Report

ABB EL Service Italy
Via Friuli 4
Dalmine
BG
24044
Italy
Phone: +396952784
Email: elserviceops@abb.com

Prepared By: Tomas Kozel
Date: 30/07/2025
CSR#: 1698295736

Customer Information

Order Address

Demo site
Street 123
54321 Dalmine
Italy

Delivery Address

Same as order address

Work Details

Work Order Number
00433323

Subject
4202815814 Secondary substation 221 ACB Maintenance and inspection

Contact Person on Site
Maintenance Manager

Cust PO Number
cash sale

Service Contract Type
Reimbursable Contract

Service SubCategory
Preventive Maintenance

CS Service Order Number

SD Sales Order Number
4202815814

Work Order Line Items

Serial Number	WOLI Number	Subject	Status

Timesheets

Date	Start Time	End Time	Hours	Category	Subject	WOLI Number	Service Resource User
28/07/2025	07:00	09:00	02:00	Preparation Time	-Prep Time	00000001	
Prep JSEA, Documents, Test equipment. Called customer.							
29/07/2025	08:00	16:36	08:36	Normal Time	-Site work	00000001	
Travel to site, Obtain security clearance and site access , Review work scope, JSEA , PTW. LOTO, carry out site work							
30/07/2025	08:00	16:30	08:30	Normal Time	-Site work, report	00000001	
Continue site work, travel back Prepare Service report							

Total 19:06

Recommendations / Additional Comments - Service Appointment

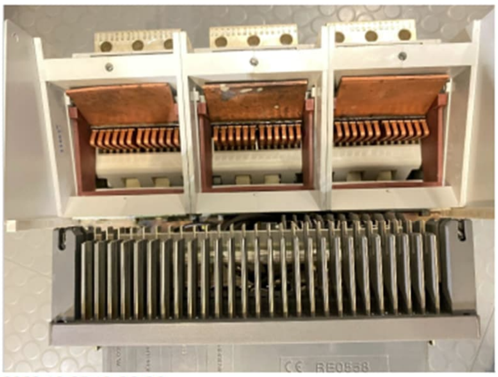
Scope:
Inspection and maintenance of two Emax ACBs, serial numbers AM51118470, AM51118458, customer tags N101-03-CB and E100-01-CB

29/07/2025
Work done on N101-03-CB (located in N101-03 panel):
Measured main contact resistance before service L1/L2/L3 = 10/10/11 MicroOhms
Minor arcing observed on power contacts (normal operational wear) - servicing contacts.
Ductor resistance after service L1/L2/L3 = 9/10/9 MicroOhms (improvement achieved)
Mechanisms require servicing because the grease is dried and there is no lubrication on the tripping mechanism - mechanism cleaned and greased for reliable operation.
Mechanical and protection relay tests satisfactory.
The N101-03-CB ACB is fit for operation.

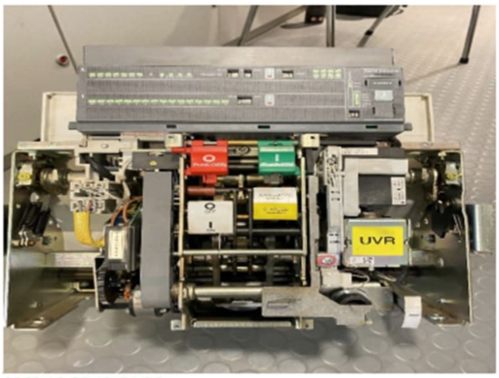
30/07/2025
Work done on E100-01-CB (located in E100-01):
Measured main contact resistance before service L1/L2/L3 = 10/10/12 MicroOhms
Minor arcing observed on power contacts (normal operational wear) - servicing contacts.
Ductor resistance after service L1/L2/L3 = 10/9/9 MicroOhms (improvement achieved)
Mechanisms require servicing because the grease is dried and there is no lubrication on the tripping mechanism - mechanism cleaned and greased for reliable operation.
Mechanical and protection relay tests satisfactory.
The E100-01-CB ACB is fit for operation.

Additional observation, risks and recommendations:
The lifting trolley for ACBs servicing is not available – risk of not being able to perform full ACB maintenance in the future – recommended to purchase a lifting trolley.

Service Report - Images



2023.10.25.10.08.00-



2023.10.25.10.01.00-

2023.10.25.09.56.00-

2023.10.25.09.56.00-_3

APPENDIX B

Structured Output of the Data Mining Process for the Sample Report

Equipment View: CSR# 1698295736

← Entity view ← Tracker

File:	*Demo_Site_Service_Report-gpt-5-1-sr-v2-9-eq-job-1772178426-420ac16d.json*	**Location:**	Demo site	**Date:**	2025-07-30

Equipment

id	serial_number	type	status	product_variant_id
eq-1	AM51118470	Emax ACB N101-03-CB (located in N101-03 panel)	Operational	

Observations

id	timestamp	type	observed_property	value	value_unit	classification
obs-1		Numerical	Measured main contact resistance before service L1	10	MicroOhms	Normal
obs-2		Numerical	Measured main contact resistance before service L2	10	MicroOhms	Normal
obs-3		Numerical	Measured main contact resistance before service L3	11	MicroOhms	Normal
obs-4		Textual	Minor arcing observed on power contacts	normal operational wear		Normal
obs-5		Numerical	Ductor resistance after service L1	9	MicroOhms	Optimal
obs-6		Numerical	Ductor resistance after service L2	10	MicroOhms	Optimal
obs-7		Numerical	Ductor resistance after service L3	9	MicroOhms	Optimal
obs-8		Textual	Mechanisms require servicing because the grease is dried and there is no lubrication on the tripping mechanism	mechanism cleaned and greased for reliable operation		Severe
obs-9		Textual	Mechanical and protection relay tests	satisfactory		Normal
obs-10		Textual	The N101-03-CB ACB	fit for operation		Normal

Risks

id	title	description	level	category
risk-1	Unavailable lifting trolley for ACB maintenance	The lifting trolley for ACBs servicing is not available – risk of not being able to perform full ACB maintenance in the future.	Medium	Reliability

Recommendations

id	title	description	recommendation_type
rec-1	Purchase lifting trolley for ACB servicing	Recommended to purchase a lifting trolley for ACBs servicing to enable full ACB maintenance in the future.	Maintenance

Actions

id	title	description	status	scheduled_date
act-1	Maintenance N101-03-CB Emax ACB	Inspection and maintenance of Emax ACB N101-03-CB (located in N101-03 panel) including contact resistance measurements, servicing contacts, mechanism cleaning and greasing, and mechanical and protection relay tests.	Completed	

Steps

id	action_reference	description	status
step-1	act-1	N101-03-CB - Measured main contact resistance before service L1/L2/L3 = 10/10/11 MicroOhms	Done
step-2	act-1	N101-03-CB - Minor arcing observed on power contacts (normal operational wear) - servicing contacts	Done
step-3	act-1	N101-03-CB - Ductor resistance after service L1/L2/L3 = 9/10/9 MicroOhms (improvement achieved)	Done
step-4	act-1	N101-03-CB - Mechanisms require servicing because the grease is dried and there is no lubrication on the tripping mechanism - mechanism cleaned and greased for reliable operation	Done
step-5	act-1	N101-03-CB - Mechanical and protection relay tests satisfactory	Done
step-6	act-1	N101-03-CB - The ACB is fit for operation	Done

Equipment

id	serial_number	type	status	product_variant_id
eq-2	AM51118458	Emax ACB E100-01-CB (located in E100-01)	Operational	

Observations

id	timestamp	type	observed_property	value	value_unit	classification
obs-11		Numerical	Measured main contact resistance before service L1	10	MicroOhms	Normal
obs-12		Numerical	Measured main contact resistance before service L2	10	MicroOhms	Normal
obs-13		Numerical	Measured main contact resistance before service L3	12	MicroOhms	Normal
obs-14		Textual	Minor arcing observed on power contacts	normal operational wear		Normal
obs-15		Numerical	Ductor resistance after service L1	10	MicroOhms	Optimal
obs-16		Numerical	Ductor resistance after service L2	9	MicroOhms	Optimal
obs-17		Numerical	Ductor resistance after service L3	9	MicroOhms	Optimal
obs-18		Textual	Mechanisms require servicing because the grease is dried and there is no lubrication on the tripping mechanism	mechanism cleaned and greased for reliable operation		Severe
obs-19		Textual	Mechanical and protection relay tests	satisfactory		Normal
obs-20		Textual	The E100-01-CB ACB	fit for operation		Normal

Risks

id	title	description	level	category
risk-1	Unavailable lifting trolley for ACB maintenance	The lifting trolley for ACBs servicing is not available – risk of not being able to perform full ACB maintenance in the future.	Medium	Reliability

Recommendations

id	title	description	recommendation_type
rec-1	Purchase lifting trolley for ACB servicing	Recommended to purchase a lifting trolley for ACBs servicing to enable full ACB maintenance in the future.	Maintenance

Actions

id	title	description	status	scheduled_date
act-2	Maintenance E100-01-CB Emax ACB	Inspection and maintenance of Emax ACB E100-01-CB (located in E100-01) including contact resistance measurements, servicing contacts, mechanism cleaning and greasing, and mechanical and protection relay tests.	Completed	

Steps

id	action_reference	description	status
step-7	act-2	E100-01-CB - Measured main contact resistance before service L1/L2/L3 = 10/10/12 MicroOhms	Done
step-8	act-2	E100-01-CB - Minor arcing observed on power contacts (normal operational wear) - servicing contacts	Done
step-9	act-2	E100-01-CB - Ductor resistance after service L1/L2/L3 = 10/9/9 MicroOhms (improvement achieved)	Done
step-10	act-2	E100-01-CB - Mechanisms require servicing because the grease is dried and there is no lubrication on the tripping mechanism - mechanism cleaned and greased for reliable operation	Done
step-11	act-2	E100-01-CB - Mechanical and protection relay tests satisfactory	Done
step-12	act-2	E100-01-CB - The ACB is fit for operation	Done

Equipment

id	serial_number	type	status	product_variant_id
eq-3		Lifting trolley for ACBs servicing	Unknown	

Observations

id	timestamp	type	observed_property	value	value_unit	classification
obs-21		Textual	The lifting trolley for ACBs servicing	not available		Severe

Failure Modes

id	description	related_observation_ids	related_failure_cause_ids
fm-1	Inability to perform full ACB maintenance due to missing lifting trolley	obs-21	fc-1

Failure Causes

id	description	related_observation_ids
fc-1	Lifting trolley for ACBs servicing not procured or not available on site	obs-21

Risks

id	title	description	level	category
risk-1	Unavailable lifting trolley for ACB maintenance	The lifting trolley for ACBs servicing is not available – risk of not being able to perform full ACB maintenance in the future.	Medium	Reliability

Recommendations

id	title	description	recommendation_type
rec-1	Purchase lifting trolley for ACB servicing	Recommended to purchase a lifting trolley for ACBs servicing to enable full ACB maintenance in the future.	Maintenance