



TECHNISCHE ANFORDERUNGEN

Revision 0.1

ZUSAMMENFASSUNG

Technische Beschreibung der Anforderungen zur Implementierung einer Digital-Process-Twin Architektur

Impressum

Website
Contact

aisoproject.com
hello@aisoproject.com



Internal Reference

Deliverable No.	D2.2
Deliverable Name	Technische Anforderungen
Lead Participant	logarithmo GmbH & Co. KG
Work Package No.	2
Task No. & Name	2.2.1
Document (File)	500-Deliverables-WP2_D2_2
Issue (Save) Date	07.10.2024

Document status

	Date	Person(s)	Organisation
Authors	01.12.2024	Tobias Quabeck	logarithmo GmbH & Co. KG
Verification by	10.12.2024	-	ZEDO, ASEW, WWNetz
Approval by			

Document sensitivity

- | | | |
|-------------------------------------|--|---|
| <input checked="" type="checkbox"/> | Not Sensitive | Contains only factual or background information; contains no new or additional analysis, recommendations or policy-relevant statements |
| <input type="checkbox"/> | Moderately Sensitive | Contains some analysis or interpretation of results; contains no recommendations or policy-relevant statements |
| <input type="checkbox"/> | Sensitive | Contains analysis or interpretation of results with policy-relevance and/or recommendations or policy-relevant statements |
| <input type="checkbox"/> | Highly Sensitive
Confidential | Contains significant analysis or interpretation of results with major policy-relevance or implications, contains extensive recommendations or policy-relevant statements, and/or contain policy-prescriptive statements. This sensitivity requires SB decision. |

Disclaimer

Der Inhalt und die Ansichten, die in diesem Material beschreiben werden, sind die der Autoren und spiegeln nicht unbedingt die Ansichten oder Meinungen der ERA-Net SES-Initiative wider. Jeder Verweis bedeutet nicht notwendigerweise die Billigung durch das ERA-Net SES.

Über ERA-Net Smart Energy Systems

ERA-Net Smart Energy Systems (ERA-Net SES) ist eine transnationale gemeinsame Plattform von 30 nationalen und regionalen Fördermittelgebern zur Initiierung von Co-Creation und Förderung von Energiesysteminnovationen. Das Netzwerk von Eigentümern und Managern nationaler und regionaler öffentlicher Förderprogramme entlang der Innovationskette bietet eine nachhaltige und dienstleistungsorientierte gemeinsame Plattform zur Förderung von Projekten in Themenbereichen wie intelligente Stromnetze, regionale und lokale Energiesysteme, Wärme- und Kältenetze, digitale Energie und intelligente Dienstleistungen usw.

Wir arbeiten mit Partnern zusammen, die helfen, die Bedürfnisse der relevanten Interessengruppen zu verstehen, und arbeiten mit Vermittlern zusammen, um ein Innovations-Ökosystem zu schaffen, das Konsortien für Forschung, Innovation, technische Entwicklung, Pilotprojekte und Demonstrationsaktivitäten unterstützt. Diese Kooperationen ebnet den Weg zur Umsetzung in realen Umgebungen und zur Markteinführung.

Darüber hinaus bietet das ERA-Net SES eine Wissensgemeinschaft, an der wichtige Demoprojekte und Experten aus ganz Europa beteiligt sind, um das Lernen zwischen Projekten und Programmen von der lokalen bis zur europäischen Ebene zu erleichtern.

www.eranet-smartenergysystems.eu

Inhaltsverzeichnis

1	Technische Anforderung für die Implementierung eines Digital Process Twin	4
1.1	Ziel und Aufbau des Berichts	4
1.2	Technische Anforderung an Datensätze.....	4
1.3	Technische Anforderung an die Schnittstellen der Quellsysteme	5
1.4	Technische Anforderung an die Integration der Daten in die Data Lake Architektur eines DPT.....	5
1.5	Technische Anforderung an den Analyse-Layer.....	6
1.6	Umsetzung im Rahmen des AISOP-Forschungsprojektes.....	6

1 Technische Anforderung für die Implementierung eines Digital Process Twin

1.1 Ziel und Aufbau des Berichts

In diesem Bericht werden die technischen Anforderungen für die Architektur eines Digital Process Twins (DPT) und ihrer Umsetzung im AISOP Projekt beschrieben. Das zentrale Ziel des Berichtes ist es a) aus der erstellten Definition des Konzeptes eines DPT eine allgemeine Liste mit technischen Anforderungen zu erstellen und b) die Implementierung der DPT-Architektur im Rahmen des AISOP Forschungsprojektes mit dem Praxispartner logarithmo GmbH & Co. KG (logarithmo) und Westfalen Weser Netz GmbH (WWN) mit Ihren Schnittstellen zu beschreiben.

Die Definition für das Konzept des DPT erfordert für die Architektur neben einem zentralen und gemeinsamen Data Lake für verschiedene Anwendungen und Schnittstellen zu den ursprünglichen Quellsystemen der Daten auch Vorverarbeitungsschritte zur Aufbereitung der Daten zu einen sogenannten Gold-Standard. Die verschiedenen Anwendungen, die auf dem zentralen Data Lake beruhen, können neben deskriptiven Datenanalysen auch weitergehende explorative Verfahren sein, wie z.B. statistische Analysen oder ML-Methoden. Im Folgenden werden die technischen Anforderungen im Allgemeinen

- an die Datensätze in den Quellsystemen,
- an die Schnittstellen der Quellsysteme,
- an die Integration der Daten in die Data Lake Architektur eines DPT und
- den Analyse-Layer

beschrieben. Die Umsetzung im Rahmen des Forschungsprojektes AISOP im Zusammenspiel von den Partnern logarithmo und WWN wird im Anschluss näher dargelegt.

1.2 Technische Anforderung an Datensätze

Aus der inhaltlichen Konzeptionierung für die Umsetzung eines DPT ergeben sich die inhaltlichen Anforderungen an die Datensätze mit den zentralen Fragestellungen: Welche Daten liegen genau innerhalb eines Prozesses vor und welche dieser Daten eignen sich für den Anwendungsfall? Daraus abgeleitet kann aus technischer Perspektive ein Datensatz grundlegend in Stamm- und Bewegungsdaten differenziert werden. *Stammdaten* umfassen dabei Datensätze, welche langfristig dauerhaft, verlässliche und aktuelle Informationen umfassen. Stammdaten stellen dabei die Grundlagen von Organisationen, Prozessen, Standorten und anderen Entitäten dar und dienen somit als Basis für weitere Berechnung und Analysen. Wichtig zu beachten ist, dass es bestimmte Fälle gibt, in denen sich Stammdaten ändern können. Gründe können beispielhaft Aktualisierung von Basis Informationen, Änderungen rechtlicher Situationen oder Anpassungen an neue Technologien und Prozesse sein. Im Gegensatz dazu stehen die *Bewegungsdaten*, welche temporäre und/oder von der Zeit abhängige Daten über Ereignisse sind, wie z.B. Messwerte der Temperatur oder die Windstärke.

Neben der reinen Klassifizierung in Stamm- und Bewegungsdaten ergeben sich für die Arbeit mit den Daten folgende technischen Anforderungen, die es für jeden Datensatz einzeln zu bewerten gilt:

- Verfügbarkeit: Sind die Daten über Schnittstellen für ein Konzept des DPT erreichbar? Gibt es Firewall oder Netzwerkeinstellungen, die einen Datentransfer beeinträchtigen können?
- Integrität: Sind die Daten vertrauenswürdig, z.B. anhand ihrer Quelle? Gibt es Mechanismen, um sicherzustellen, dass Daten nicht versehentlich gelöscht oder falsch bearbeitet werden können?
- Konsistenz: Sind die Daten übergreifend konsistent?
- Aktualität der Daten: Sind die Daten aktuell und spiegeln die tatsächliche Situation dar? In welcher Frequenz werden die Daten aktualisiert (gilt für Stamm- und Bewegungsdaten).
- Datenhistorie:
 - Stammdaten: Gibt es bereits eine Versionierung, um Änderungen nachzuvollziehen?
 - Bewegungsdaten: Wie lang reicht die Historie der Daten zurück?
- Datenformat: Liegen die Daten in einem Standard-Format wie CSV, JSON oder XML vor, um eine einfache Integration und Bearbeitung zu ermöglichen?
- Datensicherheit: Sind die Daten im Quellsystem geschützt vor Änderungen durch Dritte? Gibt es Mechanismen für Authentifizierung und Autorisierung?

1.3 Technische Anforderung an die Schnittstellen der Quellsysteme

Wie im vorherigen Abschnitt beschrieben, ist eine zentrale Anforderung für die Datensätze die Frage der Verfügbarkeit über Schnittstellen. Sofern die Frage der Übertragung geklärt ist, gilt es die passende Schnittstelle auszuwählen. Dabei gilt es zu beachten, dass eine Schnittstelle ganz allgemein in der Lage ist, Daten in einem Datenstrom zu senden und/oder zu empfangen, als auch in externe Systeme zu integrieren. Die Wahl eines Schnittstellentyps ist im Nachgang im Einzelfall zu treffen und gemäß der Vor- und Nachteile abzuwägen. Die Anforderungen an den Schnittstellentyp ergeben sich dazu aus i) der Komplexität für die Wartung von Schnittstellen aus Entwicklerperspektive, ii) der Anforderung an die Performance von Endpunkten und iii) de tn - ggf. veränderlichen - Datenstrukturen für den expliziten Anwendungsfall. Weit verbreitete Standards für eine Schnittstelle sind zum Beispiel MQTT-Broker, Rest-API oder WebSocket-Verbindungen.

1.4 Technische Anforderung an die Integration der Daten in die Data Lake Architektur eines DPT

Um eine zuverlässige und effiziente Datenintegration in die Data Lake Architektur des DPT sicherzustellen, sind bestimmte technische Anforderungen zu erfüllen:

- Datenqualität: Die integrierten Daten müssen einem „Gold-Standard“ entsprechen, was durch geeignete Data-Preprocessing-Schritte sicherzustellen ist. Als Bewertungskriterium der Datenqualität können Metriken (u.a. Vollständigkeit oder Vollständigkeit der Daten mit automatisierter Ausreißer-Analyse) herangezogen werden.
- Skalierbarkeit: Die Architektur des DPT muss skalierbar sein, um große Datenmengen verarbeiten und speichern zu können.
- Datensicherheit: Die integrierten Daten müssen sicher und zuverlässig gespeichert werden, um eine Vertrauenswürdigkeit der Ergebnisse zu gewährleisten.
- Backup-System: Um die Verfügbarkeit und Zuverlässigkeit der Datenbank zu gewährleisten, ist ein Backup Mechanismus erforderlich. Dieser Mechanismus sorgt dafür, dass eine redundante Kopie der Daten auf einem oder mehreren sekundären Systemen gespeichert wird. Wenn das primäre System ausfällt, kann die Datenbank über das redundante System weiterbetrieben werden.

1.5 Technische Anforderung an den Analyse-Layer

Um die Daten in der Data Lake Architektur des effizient analysieren und interpretieren zu können, gilt es bestimmte technische Anforderungen an den Analyse-Layer zu erfüllen. Ein wichtiger Aspekt dieser Anforderungen ist die Fähigkeit, Daten visuell darzustellen und weiterführende explorative Analysen durchzuführen. Im Detail bedeutet dies für *deskriptive Analysen* Daten visuell darzustellen, um das Verständnis für die Daten zu entwickeln. Dazu können verschiedene Darstellungsformen wie Histogramme, Boxplots oder Heatmaps gehören. *Explorative Analysen* umfassen hingegen komplexere Methoden wie Regression, Korrelation und Clusteranalyse, die es ermöglichen, Muster und Beziehungen in den Daten zu erkennen. Die weiterführenden Berechnungen sollten dabei modular gestaltet werden und einfach per Schnittstelle integrierbar sein. Technisch könnte dies z.B. über Docker-Container oder per REST-API umgesetzt werden, sodass die Analyse-Zwecke sowohl direkt in der Plattform oder extern betrieben werden können.

1.6 Umsetzung im Rahmen des AISOP-Forschungsprojektes

Im nun nachfolgenden Abschnitt werden die vorherigen Anforderungen berücksichtigt und in der Implementierung des DPT mit der zentralen Data Lake Architektur im Rahmen des AISOP-Forschungsprojektes beschrieben und eingeordnet. Mit dem formulierten Anwendungsfall Anomalien in der Betriebsplanung anhand der Verbrauchs-/Einspeiseleistung von Ortsnetzstationen (ONS) bei einem Verteilnetzbetreiber zu identifizieren, sind von dem Praxispartner WWN initial verschiedene einmalige Exporte verschiedener Datenbanksysteme geteilt worden. Diese umfassten neben Netzmodellen, EEG-Datenbankexporten auch verschiedene Messwerte auf Mittelspannungs- und Niederspannungsebene.

Für die nachfolgende Bewertung der Datensätze hat der Praxispartner logarithmo Adapter für die einzelnen Datensätze implementiert und auf die technischen Anforderungen an die Datensätze überprüft. Hierbei hat sich herausgestellt, dass die Netzdaten und Leitungsparameter zwar aus inhaltlicher Perspektive von Interesse für den Anwendungsfall sind, jedoch nicht über eine ausreichende Historie vorliegen. Dementgegen sind jedoch Messwerte von Ortsnetzstationen direkt über ein installiertes Messsystem bei der WWN passend, da diese in ausreichender Historie (über 10 Jahre) vorliegen und somit den Anforderungen besser entsprechen. Neben den von WWN geteilten Datensätzen sind für den inhaltlichen Anwendungsfall des DPT auch weitere externe Datenquellen wie Wetterdaten von Interesse und sind ausgewählt worden.

Die Integration der Daten in die DPT-Architektur und den zentralen Data Lake ist über standardisierte REST-API Schnittstellen umgesetzt. Aufgrund der Flexibilität die Daten in unterschiedlichen Formaten zu empfangen, der guten Skalierbarkeit und der einfachen Integration ist die Wahl auf REST-APIs gefallen. Dies umfasst den kontinuierlichen Austausch von Messwerten an ausgewählten ONS im Gebiet von WWN zu logarithmo (über ONS mit ausgestattetem LoRaWAN-Sensoren) wie auch die Schnittstelle zu den DWD-Wetterdaten, welche gezielt für die untersuchten Netzregionen im WWN-Netz ausgewählt worden sind. Unter Berücksichtigung der technischen Anforderungen an die Integration der einzelnen Datensätze in die Data Lake Architektur werden die Datensätze für den anschließende Analyse-Layer aufbereitet und können sowohl in Form einer Schnittstelle direkt aus dem Data Lake verfügbar gemacht werden, als auch innerhalb des DPT in Form von einer eigenen Web-Applikation abgerufen und visuell aufbereitet werden.

FUNDING



This project has received funding in the framework of the joint programming initiative ERA-Net Smart Energy Systems' focus initiative Digital Transformation for the Energy Transition, with support from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883973.