

Reihe 20

Rechnerunter-
stützte Verfahren

Nr. 476

M.Sc. Christian von Trotha,
Aachen

Entwicklung eines Konzepts zur Kontextualisierung von Prozessdaten

ACPLT
AACHENER
PROZESSLEITECHNIK

Lehrstuhl für
Prozessleittechnik
der RWTH Aachen

Entwicklung eines Konzepts zur Kontextualisierung von Prozessdaten

Von der Fakultät für Georessourcen und Materialtechnik der
Rheinisch-Westfälischen Technischen Hochschule Aachen

zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

genehmigte Dissertation

vorgelegt von

M.Sc. Christian von Trotha

aus Offenbach am Main

Berichter: Univ.-Prof. Dr.-Ing. Ulrich Epple
Univ.-Prof. Dr.-Ing. Christian Diedrich
Univ.-Prof. Dr.-Ing. Tobias Kleinert

Tag der mündlichen Prüfung: 23.04.2021

Fortschritt-Berichte VDI

Reihe 20

Rechnerunterstützte
Verfahren

M.Sc. Christian von Trotha,
Aachen

Nr. 476

Entwicklung eines
Konzepts zur
Kontextualisierung
von Prozessdaten

von Trotha, Christian

Entwicklung eines Konzepts zur Kontextualisierung von Prozessdaten

Fortschr.-Ber. VDI Reihe 20 Nr. 476. Düsseldorf: VDI Verlag 2021.

188 Seiten, 82 Bilder, 12 Tabellen.

ISBN 978-3-18-347620-6, ISSN 0178-9473,

€ 67,00/VDI-Mitgliederpreis € 60,30.

Für die Dokumentation: Kontextualisierung – Datenbasierte Entscheidung – Wissensgraph – Vernetzung von Informationen – Prozessdaten – Assistenzsystem – Graphdatenbank

Keywords: Contextualization – data-based decision – knowledge graph – interlinked information – process data – operator support system – graph database

Die vorliegende Arbeit wendet sich an Ingenieur:innen und Wissenschaftler:innen aus der Prozessindustrie. Sie befasst sich mit der Vernetzung und der Verknüpfung von verschiedenen Prozessdaten. Grundlage dafür sind die in diversen technischen Dokumenten, wie zum Beispiel dem R&I-Diagramm, verfügbaren Strukturinformationen. Dank der Vernetzung bzw. der Verknüpfung können die verfügbaren Prozessdaten deutlich effizienter gefiltert und durchsucht werden. Dadurch kann der situationsbezogene Kontext bereitgestellt werden und als Entscheidungsgrundlage dienen. Kern der Arbeit ist eine Modellhierarchie, die die Modellierung der Abhängigkeiten in einem Wissensgraphen ermöglicht. Demonstriert wird das Konzept anhand der Laboranlage des Lehrstuhls. Grundlage des Demonstrators ist die Graphdatenbankplattform Grakn. Verschiedene Verwaltungsschalen, BaSys40-Komponenten und PandIX dienen dabei als Informationsquellen.

Bibliographische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet unter www.dnb.de abrufbar.

Bibliographic information published by the Deutsche Bibliothek

(German National Library)

The Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie (German National Bibliography); detailed bibliographic data is available via Internet at www.dnb.de.

D82 (Diss. RWTH Aachen University, 2021)

© VDI Verlag GmbH · Düsseldorf 2021

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe (Fotokopie, Mikrokopie), der Speicherung in Datenverarbeitungsanlagen, im Internet und das der Übersetzung, vorbehalten.

Als Manuskript gedruckt. Printed in Germany.

ISSN 0178-9473

ISBN 978-3-18-347620-6

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit am Lehrstuhl für Prozesstechnik der RWTH Aachen. An dieser Stelle möchte ich mich bei allen bedanken, die mir geholfen haben, sie erfolgreich abzuschließen.

Besonderer Dank gilt Herrn Prof. Dr.-Ing. Ulrich Epple für die ermöglichten Freiräume, die konstruktiven Diskussionen und die förderliche Arbeitsatmosphäre am Lehrstuhl. Sowohl der intensive fachliche Austausch als auch das kreative Arbeitsumfeld stellen die Basis für den erfolgreichen Abschluss dieser Arbeit dar.

Herrn Prof. Dr.-Ing. Christian Diedrich danke ich für die Übernahme des Koreferats und Herrn Prof. Dr.-Ing. Pfeifer für die Übernahme des Prüfungsvorsitzes.

Herrn Prof. Dr.-Ing. Tobias Kleinert danke ich für die konstruktiven Hinweise zu meiner Arbeit und wünsche viel Erfolg bei der Fortführung des Lehrstuhls.

Meinen ehemaligen Kolleg:innen danke ich für die anregenden Diskussionen, den intensiven Austausch und die angenehme sowie produktive Arbeitsatmosphäre. Ganz besonderer Dank gilt Frau Margarete Milesu für die Unterstützung bei der Navigation durch die vielen administrativen Hürden des Unialltags.

Bei meiner Frau Almut möchte ich mich von Herzen für die unendliche Geduld und Unterstützung bedanken. Außerdem möchte ich mich bei meinem Freund Markus bedanken, der mir immer mit Rat und Tat zur Seite stand.

Abschließend möchte ich mich ganz besonders bei meinen Eltern Sabine und Eckhard und meinen Schwiegereltern Grete und Peter bedanken. Ohne die liebevolle und großzügige Unterstützung meiner Eltern wäre diese Arbeit nicht möglich und ohne die herzliche und kulinarische Unterstützung meiner Schwiegereltern nur halb so angenehm gewesen.

Vielen Dank!

Aachen, im Mai 2021

Christian v. Trotha

Inhaltsverzeichnis

Abkürzungen und Symbole	VIII
Kurzfassung	X
Abstract	XI
1 Einleitung	1
1.1 Motivation	1
1.2 Vorgehensweise	4
1.2.1 Methoden zur Ontologieentwicklung aus der Literatur	4
1.2.2 Vorgehensweise und Struktur der Arbeit	6
2 Definition von Zielszenarien	7
2.1 Verknüpfung von Informationen zu Wissen	7
2.2 Aufdeckung von bisher nicht modellierten Zusammenhängen	7
2.3 Erstellung von Datensubsets zur Datenanalyse	8
2.4 Validierung von Ergebnissen einer Datenanalyse	9
2.5 Analyse der Zielszenarien	9
3 Grundlagen	10
3.1 Eigene Vorarbeiten	10
3.1.1 Assistenzsysteme	10
3.1.2 Strukturmodellierung	11
3.1.3 Steuerungsarchitektur	11
3.2 Datenqualität	13
3.3 Modelle und Modellbildung	14
3.4 Wissensrepräsentation	18
3.4.1 Semantische Netze	22
3.4.2 Graphdatenbanken	23
3.5 Definition des Begriffs Kontext für die Prozessindustrie	25
3.6 Ursache-Wirkungsanalyse	28
3.7 Verwandte Ansätze aus der Literatur	29
4 Bestehende Modellierungskonzepte und Informationsquellen	33
4.1 Systeme	33
4.1.1 Systemmodellierung	34
4.1.2 Rollenmodell	35
4.1.3 Konzeptuelles Datenmodell der ISO 15926	35
4.1.4 Grakn-Metamodell	37

4.2	Datenquellen	37
4.2.1	Strukturmetamodelle	37
4.2.2	Verwaltungsschale	44
4.2.3	Prozesshistorian	45
4.3	Ereignisse	46
4.3.1	Abnormale Situationen	46
4.3.2	Ereignistypen	47
4.4	Diskussion der Informationsquellen	48
5	Anforderungsanalyse und -spezifikation	49
5.1	Analyse der Randbedingungen aus dem Anwendungsgebiet	49
5.1.1	Klassische Randbedingungen der Prozessindustrie	50
5.1.2	Neue Randbedingungen durch Industrie 4.0	51
5.2	Anforderungen an das Konzept	51
5.3	Anforderungen an die Modellhierarchie	54
5.4	Anforderungen des Konzepts an die Umwelt	55
6	Entwicklung der Systemstruktur	57
6.1	Definition der Systemarchitektur	57
6.2	Entwicklung der Modellhierarchie	59
6.3	Beziehungsmodell	59
6.3.1	Konzept	59
6.3.2	Anwendungsbereich	60
6.3.3	Allgemeine Spezifikation	60
6.3.4	Formale Spezifikation	60
6.3.5	Diskussion des Beziehungsmodells	62
6.4	Spezifikation des Metametamodells	62
6.4.1	Objekt (object)	63
6.4.2	Subjekt (subject)	63
6.4.3	Eigenschaften (property)	63
6.4.4	Lebenszyklus (lifecycle)	64
6.4.5	Ereignis (event)	65
6.4.6	Kontextbeziehung (context relation)	65
6.4.7	Entität (entity)	66
6.4.8	Rollen (role)	67
6.4.9	Anwendung des Metametamodells	67
6.5	Aspektmetamodelle	68
6.5.1	Aspektmetamodell Topologie	69
6.5.2	Aspektmetamodell Messen	73
6.5.3	Aspektmetamodell Deployment	75
6.5.4	Aspektmetamodell Steuerungsstruktur	77
6.5.5	Aspektmetamodell Hardware-Software-Kopplung	80
6.6	Mögliche Erweiterungen zur Vereinfachung des Datenzugriffs	81
6.6.1	Definition von Perspektiven	81
6.6.2	Analyse der Wirkrichtung der Beziehungen	81
6.7	Definition eines Prozesses zur modularen Erweiterung	83

6.8	Evaluation der Modellhierarchie	84
6.8.1	Evaluation anhand der Anforderungen an die Modellhierarchie	84
6.8.2	Evaluation anhand der Anforderungen an das Konzept	85
6.8.3	Klassifikation als Assistenzsystem	86
7	Realisierung und prototypische Implementierung	88
7.1	Demonstration am Pumpwerk	88
7.1.1	Datenquelle: openAAS-Server	89
7.1.2	Datenquelle: Prozesshistorian	91
7.1.3	Datenquelle: Einzelsteuereinheiten	91
7.1.4	Datenquelle: Engineering	92
7.2	Verwendete Werkzeuge	92
7.2.1	GRAKN.AI	92
7.2.2	ACPLT/RTE	95
7.3	Modellierung	97
7.3.1	Abbildung des Metametamodells auf das Grakn-Metamodell	97
7.3.2	Abbildung der Aspektmetamodelle auf das Grakn-Metamodell	102
7.3.3	Abbildung der Datenquellen auf die Aspektmetamodelle	102
8	Evaluation der Ergebnisse	103
8.1	Evaluation anhand der Anforderungen	103
8.2	Evaluation anhand der Anwendungsszenarien	105
8.2.1	Verknüpfung von Informationen zu Wissen	105
8.2.2	Aufdeckung von bisher nicht modellierten Zusammenhängen	108
8.2.3	Erstellung von Datensubsets zur Datenanalyse	110
8.2.4	Validierung von Ergebnissen der Datenanalyse	110
8.2.5	Weitere Beziehungstypen	112
9	Diskussion und Ausblick	116
A	Anhang	119
A.1	Pandix-Datei des Pumpwerks	119
A.2	Abbildung der Aspektmetamodelle in Graql	143
A.2.1	Aspektmetamodell Topologie	143
A.2.2	Aspektmetamodell Messen	145
A.2.3	Aspektmetamodell Deployment	146
A.2.4	Aspektmetamodell Steuerungsstruktur	146
A.2.5	Aspektmetamodell Hardware-Software-Kopplung	147
A.2.6	Aspektmetamodell Wirkung	148
A.3	GraphML-Dateien der Nachbarschaft von L26	148
Literaturverzeichnis		166

Abkürzungen und Symbole

- ACPLT/RTE** Aachener Prozessleittechnik/ Run Time Environment
- ASM** Abnormal Situation Management
- CAE** Computer Aided Engineering
- CAEX** Computer Aided Engineering eXchange
- CPS** Cyber-Physisches System
- DDL** Data Definition Language
- DEXPI** Data EXchange in the Process Industry
- DML** Data Manipulation Language
- DOM** Document Object Model
- DYMASOS** DYnamic MAnagment of physically coupled System of Systems
- ER** Entity-Relationship
- ESKAPE** Evolving Sematic Knowledge Aggregation and Processing Engine
- JSON** JavaScript Object Notation
- LCA** Lebenszykluseintragsarchiv (LiveCycle Entry Archive)
- LC_SM** Lebenszyklus-Teilmodell (LifeCycle SubModel)
- M2M** Maschine-zu-Maschine-Kommunikation
- OMG** Object Management Group
- OPC UA** Open Platform Communication Unified Architecture
- OWL** Ontology Web Language
- PandIX** Pipe and Instrumentation eXchange
- PLT** Prozessleittechnik
- PPE** ProcessPlantElement
- RDF** Resource Description Framework
- R&I** Rohrleitungs- und Instrumentenfließschema
- RIO** Remote Input/Output
- RIVA** Rohrleitungs- und Instrumentenfließschema verfahrenstechnisch aufbereitet
- SAX** Simple API for XML
- SIC** Systemelement Interface Connection
- SoSX** System of System eXchange
- SPS** speicherprogrammierbare Steuerung

- UML** Unified Modeling Language
URI Uniform Resource Identifier
W3C World Wide Web Consortium
XML Extensible Markup Language
XSD XML Schema Definition

Kurzfassung

In der Prozessindustrie werden im Rahmen des Betriebs prozesstechnischer Anlagen mit zunehmendem Umfang Daten erzeugt und abgespeichert. Dabei handelt es sich beispielsweise um Daten aus Planungsprozessen (z. B. das R&I-Fließbild) oder um Prozessdaten (z. B. Messreihen). Für den Austausch und Zugriff auf diese Daten setzen sich standardisierte Schnittstellen und Austauschformate, wie zum Beispiel die Verwaltungsschale, OPC UA, DEXPI oder PandIX immer stärker durch und ermöglichen so eine eindeutige Identifikation und semantische Referenzierung der abgelegten Daten. Darauf aufbauend ist der nächste Schritt auf dem Weg zu datenbasierten Entscheidungen die Aufbereitung und Kontextualisierung dieser Daten. Diese Kontextualisierung und Aufbereitung muss aufwandsarm und mithilfe von standardisierten Werkzeugen erfolgen, damit die Ausnutzung des bestehenden Datenschatzes schnell und effizient erfolgen kann.

Die vorliegende Arbeit adressiert die Herausforderung der Kontextualisierung. Sie beschreibt ein Konzept, das die konsistente Modellierung des Kontextes von Prozessdaten in der Prozessindustrie ermöglicht. Kern der Arbeit ist eine Modellhierarchie, mithilfe derer die Zusammenhänge zwischen einzelnen Daten dargestellt werden können. Auf diesem Weg können die verfügbaren Informationen zu Wissen vernetzt werden. Dabei wird ausgenutzt, dass in der Prozessindustrie einerseits vielfältige Informationen über strukturelle Zusammenhänge, beispielsweise in Form von Schalt-, Bau- und Aufstellungsplänen, Fließbildern oder Steuerungsarchitekturen, beschrieben sind und dass sich andererseits lokale Abhängigkeiten zwischen einzelnen Daten gerade aus diesen Strukturen ergeben. So ermöglicht das Kontextualisierungssystem eine effiziente Suche und Filterung der verfügbaren Daten und erleichtert so die Nutzung dieser.

Technologische Grundlage der Arbeit ist eine Wissensrepräsentation mithilfe eines semantischen Netzes. Semantische Netze repräsentieren Informationen mithilfe von Knoten und Kanten und eignen sich insbesondere zur Beschreibung von Zusammenhängen und Abhängigkeiten zwischen Informationen. Damit bilden sie die ideale Grundlage für die Kontextualisierung. Eine wesentliche Herausforderung beim Einsatz von semantischen Netzen ist die oft fehlende formale Spezifikation der Begriffe und Beziehungen. Diese Herausforderung wird im Rahmen dieser Arbeit und für den Anwendungsfall der Kontextualisierung von Prozessdaten mithilfe der oben genannten Modellhierarchie adressiert.

Die praktische Anwendbarkeit des Konzepts wird mithilfe einer prototypischen Implementierung für das institutseigene Pumpwerk demonstriert. Sie basiert auf der Graphendatenbankplattform Grakn und baut auf verschiedenen Verwaltungsschalen, BaSys40-Komponenten sowie PandIX als Informationsquellen auf.

Abstract

Data is being generated and stored during the operation of process plants. This is, for example, data from planning processes (e.g. the P&ID flow chart) or process data (e.g. measurement series). Standardized interfaces, information models and exchange formats, such as the asset administration shell, OPC UA, DEXPI or PandIX are becoming increasingly popular for the exchange and access to this data. This enables clear identification and semantic referencing of the stored data. The next step to accomplish data-based decisions is the preparation and contextualization of this data. This contextualization and preparation must be carried out with little overhead and using standardized tools so that the existing data can be used easily and efficiently.

This work addresses the challenge of contextualization. It describes a concept that enables consistent modeling of the data context in the process industry. The core of the work is a model hierarchy, which can be used to model the relationships between individual data points. In this way, the available information can be interconnected to form knowledge. The concept takes advantage of the wide range of available information about structural relationships, for example in the form of circuit, construction and installation plans, flow diagrams or control architectures. These structures determine the local dependencies between individual data points. Thereby the contextualization system enables the efficient search and filtering of the available data, making it easier to use them.

The technological basis of the work is a knowledge representation using a semantic network. Semantic networks represent information using nodes and edges. They are particularly suitable for describing relationships and dependencies between data points and form the ideal basis for contextualization. A major challenge of using semantic networks is the lack of formal specification of terms and relationships. For the use case of contextualization of process data this challenge is addressed in this work by employing the model hierarchy mentioned above.

The applicability of the concept is demonstrated by means of a prototype implementation for the lab system “pumping station”. It is based on the Grakn graph database platform and uses various asset administration shells, BaSys40-components and PandIX as information sources.

