

Fachartikel

Anwendungsszenarien für Wissensnetze bei Siemens

Industrial Knowledge Graph

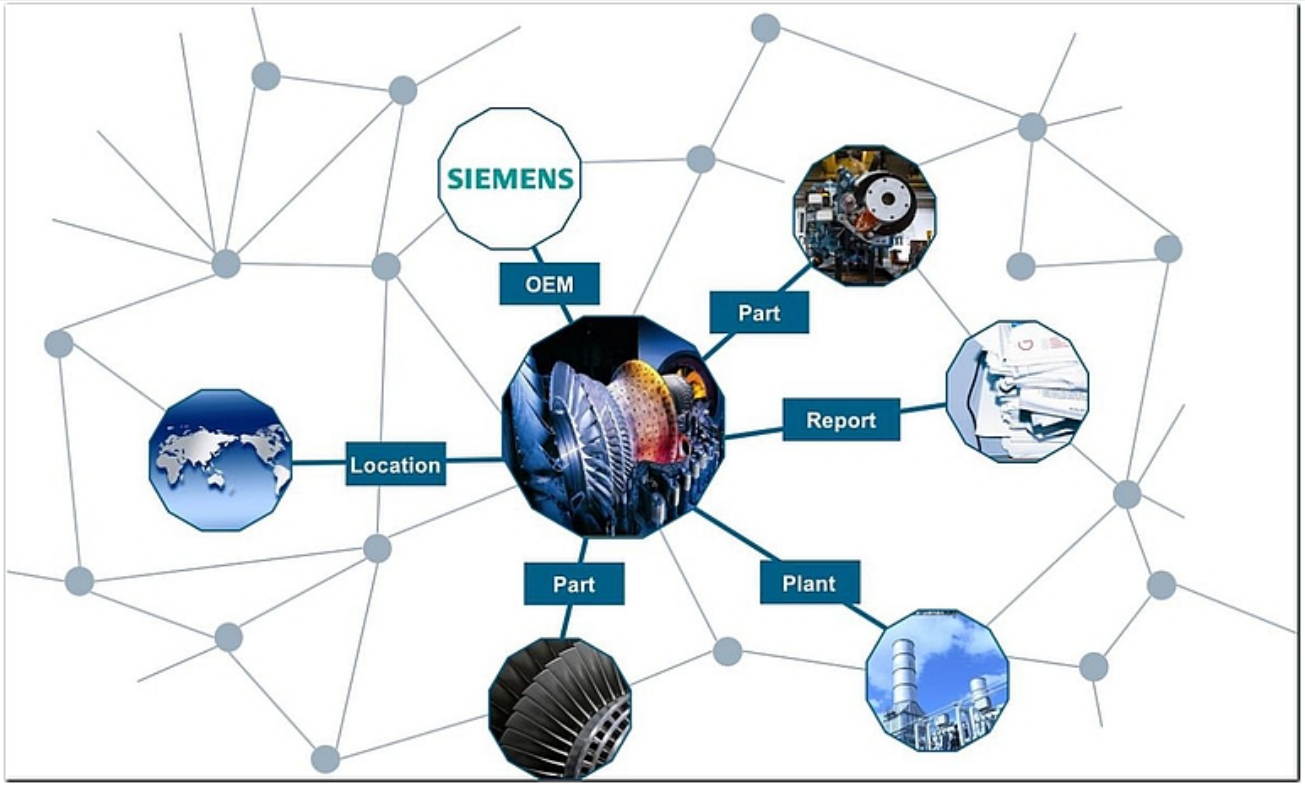
von Thomas Hubauer, Steffen Lamparter, Peter Haase, Daniel Herzig-Sommer

Der Industrial Knowledge Graph hat sich zu einem Kernelement der Siemens Digitalisierungsstrategie entwickelt. Im folgenden Artikel stellen wir dar, wie semantische Technologien bei Siemens erfolgreich in unterschiedlichsten Anwendungsfällen aus den Bereichen Engineering, Gebäudemanagement und Innovationsmanagement eingesetzt werden.

Wenn Siemens wüsste, was Siemens weiß

Wie zahlreiche andere Industrieunternehmen durchläuft auch Siemens eine umfassende Digitalisierung praktisch aller Bereiche, sowohl im Kundengeschäft als auch unternehmensintern. An vielen Stellen treffen Mitarbeiter dabei auf dieselben, wiederkehrenden Herausforderungen:

Benötigte Daten sind häufig verteilt auf mehrere *isolierte Silos*. Neben der fast klassischen Trennung der Daten nach Eigentümer (Kunde, Siemens Divisionen, ...) finden sich häufig alternativ (oder zusätzlich) Silos für bestimmte Dateitypen (z. B. Dokumente oder Zeitreihendaten) oder Informationstypen (Wartungsdaten, Stammdaten, operative Daten und andere).
Damit verbunden ist teilweise eine *geringe Datenqualität* der erfassten Informationen festzustellen. Neben der Aktualität der Daten stellen hier insbesondere Duplikate und das oft damit verbundene Problem von fehlerhaften oder widersprüchlichen Daten eine große Herausforderung für die Nutzer dar.
Darüber hinaus haben Nutzer oft mit *komplexem Datenzugriff* zu kämpfen. Ein Aspekt ist hierbei die Schwierigkeit, benötigte Daten zu finden. Neben einer fehlenden Übersicht über die unterschiedlichen Silos macht sich hier auch das Fehlen einer integrierten Suchmöglichkeit negativ bemerkbar. Doch selbst wenn die benötigten Quellen bekannt sind, wird der Zugriff auf die eigentlichen Daten oft durch hochkomplexe Datenmodelle verkompliziert, welche sich dem Domänenexperten nicht erschließen. In vielen Fällen ist es notwendig, einen IT-Experten zur Datenbereitstellung einzubinden.
Zusammengefasst führen diese Schwierigkeiten oft zu *ineffizienten Abläufen*, und damit verbunden einer unnötig langen „time-to-data“.



Der Industrial Knowledge Graph – Intelligente Informationsdrehscheibe und Quelle neuen Wissens

In den letzten Jahren hat sich der Industrial Knowledge Graph quer durch Siemens als vielversprechende Technologie etabliert. Die Technologie unterstützt die unternehmensweite Digitalisierung dabei je nach Anwendungsfall auf unterschiedliche Arten.

Ontologien und Knowledge Graphs sind Kernelemente der intelligenten Informationsdrehscheibe. Als nutzerfreundliche logische Datenmodelle ermöglichen Ontologien einen Datenzugriff ohne die Kenntnis komplexer Datenbank-Schemata. Gleichzeitig ermöglicht die explizite Semantik der Modelle es, unterschiedliche Silos durch gemeinsame Modell-Elemente zu verknüpfen und so größere Datenlandschaften zu erstellen, die sogar (teil-)automatisiert auf Konsistenz und Datenqualität geprüft werden können. Gleichzeitig ist die große Flexibilität des zugrunde liegenden Graph-Datenmodells unerlässlich zur Zusammenführung heterogener Informationen.

Den scheinbaren Widerspruch zwischen hoher Flexibilität und starker Formalisierung des Datenmodells lösen wir durch ein Schichtenmodell: flexible, schwach formalisierte Modelle werden Ebene für Ebene mit zusätzlichen Regeln und Einschränkungen angereichert, um schließlich eine automatisierte Verarbeitung zu ermöglichen.

In vielen Anwendungen ist es allerdings notwendig, Daten nicht nur in hoher Qualität und semantisch angereichert zur Verfügung zu stellen, sondern neues Wissen aus vorhandenen Informationen zu generieren. Hierfür nutzen wir Machine Learning. Während klassische Verfahren hier insbesondere von der erfolgten Informationsintegration profitieren, fokussieren wir uns vor allem auf Verfahren des Relational Machine Learning, welche unter Ausnutzung der Graphstruktur in vielen Fällen Modelle besserer Qualität liefern.

Die notwendigen Komponenten zur Realisierung dieser Funktionalitäten sind – teilweise sogar kostenlos – verfügbar. Innerhalb von Siemens gibt es unterschiedliche technologische Stacks zur Realisierung von Industrial Knowledge Graphs.

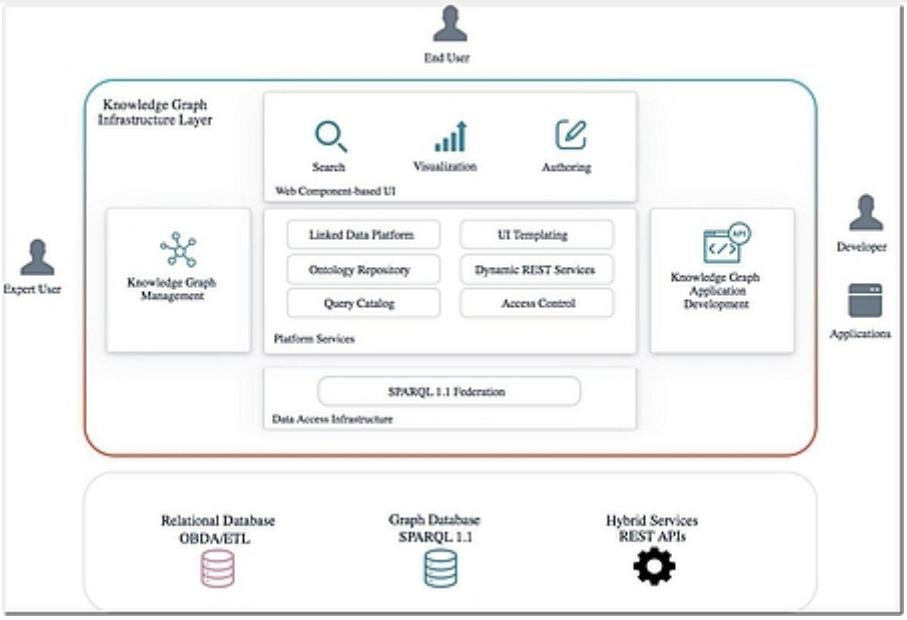


Abb. 1: RDF-basierter Stack für Knowledge Graph-Anwendungen bei Siemens

Abbildung 1 stellt einen häufig eingesetzten, RDF basierten Stack dar. Als Datenbanktechnologie kommt eine zu SPARQL 1.1 kompatible Graphdatenbank wie zum Beispiel Amazon Neptune zum Einsatz. Diese wird ergänzt durch eine Infrastruktur-Schicht, welche den Nutzern unterschiedliche Mehrwertdienste zur Verfügung stellt. Dazu zählen unter anderem konfigurierbare webbasierte Benutzeroberflächen, Zugriffskontrolle, eine Federation-Schicht zur flexiblen Anbindung externer Daten (beispielsweise über SQL- oder REST-Schnittstellen) sowie die Fähigkeit, parametrisierte Abfragen als REST-Endpunkte anderen Anwendungen zur Verfügung zu stellen.

Zur Umsetzung der im Folgenden beschriebenen Anwendungsfälle kam metaphacts' Knowledge Graph Platform – metaphactory – zum Einsatz. Neben RDF (Resource Description Framework) nutzen wir auch PropertyGraph-Technologie, insbesondere bei stark analytisch geprägten Anwendungen.

In den folgenden Abschnitten stellen wir exemplarisch drei Anwendungen dieser Technologie innerhalb Siemens vor.

Jeder ist ein Data Scientist – Turbinendaten im Direktzugriff

Service und Wartung von Gasturbinen sind für Siemens ein bedeutendes Geschäftsfeld. Bei diesen Turbinen handelt es sich teilweise um hochkomplexe Systeme mit stark spezialisierten, teilweise sogar in 3D-Druckverfahren produzierten Bauteilen. Aufgrund der daraus resultierenden hohen Produktionsdauern und -kosten ist es für Siemens von großem Interesse, Teile nicht unnötig oder verfrüht auszutauschen. Ebenso hat der Kunde ein großes Interesse daran, Ausfallzeiten der Turbine weitest möglich zu reduzieren.

Beides ist nur erreichbar, wenn für die Wartungsplanung wie auch für verwandte R&D-Aktivitäten qualitativ hochwertige Informationen verschiedenster Bereiche zugegriffen und miteinander in Beziehung gesetzt werden können. Diese Daten umfassen unter anderem:

Messungen von an und in der Turbine verbauten Sensoren,
Metainformationen betreffend Standort, Klima, und Produktgeneration,
Informationen über die einzelnen Komponenten und Bauteile inklusive Details zu eventuellen Zulieferern, sowie detaillierte Reparaturinformationen.

Diese umfangreichen (teilweise historischen, teilweise tagesaktuellen) Daten waren für den Fachexperten aufgrund der Vielzahl unterschiedlicher Datenquellen und der Komplexität der Modelle bislang nicht direkt verfügbar. Stattdessen musste der Turbinen-Experte sich im ersten Schritt die benötigten Daten durch einen IT-Experten zusammenstellen und aufbereiten lassen. Die eigentliche Analyse konnte so oft nur verzögert beginnen.

Domänen- und Technologieexperten haben in den letzten Jahren gemeinsam einen Industrial Knowledge Graph erstellt, der bereits jetzt einen Großteil der relevanten Quellen umfasst und kontinuierlich erweitert wird (aktuell beispielsweise wird die

Anbindung der Zeitreihendatenbank umgesetzt). Dieser integrierte Graph ermöglicht es den Experten, benötigte Daten selbst mittels eines grafischen Query Builders zu definieren und etablierten Analysetools wie Tableau oder Spotfire zur Verfügung zu stellen.

Je nach Datentyp wurden dabei unterschiedliche Strategien zur Integration angewandt: Wartungsdokumente beispielsweise werden durch eine automatische Natural Language Processing (NLP) Pipeline analysiert, relevante Reparatur-Ereignisse extrahiert und als Fakten in den Knowledge Graph geschrieben. Die Anbindung von Laufzeitdaten wie beispielsweise Temperaturmessungen aus der Brennkammer hingegen erfolgt virtuell, das heißt, die Daten verbleiben im Quellsystem (einer relationalen Datenbank) und werden vom Knowledge Graph erst bei Bedarf über sogenannte Mappings dort abgefragt. Für den Benutzer erfolgt dieser Zugriff allerdings komplett transparent: Er definiert grafisch eine Abfrage gegen das Datenmodell des integrierten Knowledge Graph und erhält als Ergebnis eine integrierte Antwort als Basis für seine weiterführenden Untersuchungen. Implementierungsdetails wie Daten-Lokalisation, Datenbank-Schemata und Logik zur Integration der Quellen bleiben ihm verborgen.

Analysen nach der Einführung des Systems belegen den deutlich schnelleren Informationszugriff für Anwender sowie eine signifikante Verringerung an Anfragen an das Datenbank-Team zur Informationsbereitstellung. Äußerst positives Feedback der Service-Ingenieure belegt außerdem, dass Industrial Knowledge Graphs nicht nur die Wertschöpfung in internen Prozessen verbessern, sondern zusätzlich auch die Zufriedenheit der Mitarbeiter signifikant verbessern.

Lernende Gebäude – Intelligente Mehrwertdienste in der Gebäudeautomatisierung

Moderne Gebäude sind aus unterschiedlichsten Gründen mit einer großen Menge verschiedener Sensoren ausgestattet. So überwachen und beeinflussen Temperatur- und Feuchtigkeitssensoren die Klimasteuerung, während zum Beispiel Licht- und Stromnutzung zu Abrechnungszwecken kleinteilig erfasst werden. Aus Brandschutzgründen vorgeschriebene Rauch- und Feuermelder liefern heutzutage umfangreiche Daten über die Zusammensetzung der Raumluft. Im Bereich Building Technologies entwickelt Siemens neuartige Geschäftsmodelle unter Nutzung dieser Daten.

Im Rahmen dieser Aktivitäten nutzen wir einen Industrial Knowledge Graph als Infrastruktur zur Realisierung des Building Digital Twin. Dieser führt Design-Informationen des Gebäudes (beispielsweise im IFC-Format) zusammen mit Informationen zur Aufteilung (etwa Mietverhältnisse) sowie den Zeitreihen-Daten der verbauten Sensoren. Auf Basis dieser Informationen realisiert Siemens Building Technologies verschiedene innovative Lösungen.

Der Verwalter einer Industrieimmobilie muss beispielsweise sicherstellen, dass alle gesetzlichen Vorgaben etwa hinsichtlich Brandschutz erfüllt werden. Beispiele hierfür sind Vorgaben zur Platzierung von Sensoren, welche Faktoren wie Raumgröße, Verbindung von Räumen, aber auch Faktoren wie Belüftung mit einbeziehen. Insbesondere ist es nach einer Umplanung der Raumaufteilung oder einer Änderung des Lüftungskonzepts daher notwendig, das Gebäude erneut hinsichtlich der Vorgaben zu prüfen.

Der Digital Twin eines Gebäudes automatisiert diese Prüfungen basierend auf Geometrieinformationen aus den IFC-Quellen sowie abgeleiteten Informationen wie Verbindungen zwischen Räumen. Der Zeitaufwand für Umplanungen kann so signifikant verringert werden.

Derselbe Digital Twin unterstützt auch weiterführende analytische Anwendungen. So haben wir beispielsweise eine Komponente entwickelt, die basierend auf den vorhandenen Sensordaten verlässlich den Belegungszustand eines Raums vorhersagen kann. Das grafbasierte Datenmodell ermöglicht es hier, auch die Messungen „beweglicher“ Sensoren (etwa in einer Lampe verbaut) basierend auf Positionsdaten einzelnen Räumen zuzuordnen. Basierend auf der entwickelten Komponente können komplexe Services angeboten werden, beispielsweise ein Companion, der die Aufteilung großer beziehungsweise Zusammenlegung kleiner Besprechungsräume basierend auf historischen Nutzungsinformationen vorschlägt, oder aber eine erweiterte Raumsuche, die Personen im Gebäude direkt zum nächsten freien Raum mit bestimmten Kriterien führt – wobei hier der tatsächliche Laufweg zum Raum einbezogen wird.

Lose Enden verknüpfen – Knowledge Graph im internen R&D-Management

Aktuell entwickeln wir auf Basis von Industrial Knowledge Graphs eine Plattform zur Vernetzung von Forschern und ihrer Projekte. Ziel dieser Entwicklung ist es, projekt- wie bereichsübergreifend Transparenz über Projekte und relevante Forschungsthemen zu schaffen. Besondere Bedeutung hat hier die Anbindung existierender interner Systeme beispielsweise zur Verwaltung von Projekten, von Mitarbeiterdaten oder Informationen zur Siemens Organisationsstruktur. Nur so können unnötige Mehrfacheingaben vermieden und Daten auch längerfristig aktuell und konsistent gehalten werden. Ermöglicht wird dies durch die in **Abbildung 1** dargestellte Fähigkeit unserer Infrastruktur, externe Systeme über APIs anzubinden und auch derartige Informationen virtuell in den Graph zu integrieren.

Die entwickelte Plattform gibt Projektleitern und R&D-Verantwortlichen die Möglichkeit, den Projektbestand flexibel anhand beliebiger Kriterien zu durchsuchen. Unterstützt durch Visualisierungen können so Überschneidungen (thematisch wie personell) schnell identifiziert und mögliche Synergien realisiert werden. Dies ist jedoch abhängig von der Information (etwa Zuordnung eines Projekt zu Themen), die der jeweilige Projektleiter vorgenommen hat.

Erfahrungsgemäß steht das Pflegen derartiger Informationen jedoch häufig unter Zeitdruck, während Projektbeschreibungen oft deutlich ausführlicher sind, da sie beispielsweise auch im Intranet oder in anderen Systemen wiederverwendet werden können. Um dem Rechnung zu tragen, haben wir den Projektgraphen um die Fähigkeit erweitert, dem Projektleiter basierend auf seinen beschreibenden Texten relevante Themen und Technologien zur Übernahme vorzuschlagen. Darauf aufbauend haben wir eine Komponente entwickelt, die basierend auf Gemeinsamkeiten zwischen Projekten (hinsichtlich Personen, Themen und beteiligten Einheiten) Projektleitern wie auch R&D-Managern mögliche Kollaborationen vorschlägt. Erweitert um eine Feedbackfunktion zu den Vorschlägen wird der Projektgraph in Zukunft kontinuierlich lernen, wo sinnvollerweise Kollaborationen initiiert werden sollen.

Zusammenfassung und Ausblick

Siemens befindet sich mitten in einem Prozess der digitalen Transformation. An drei Beispielen aus unterschiedlichsten Unternehmensbereichen haben wir demonstriert, wie Knowledge Graph und Machine Learning in diesem Umfeld Mehrwert schaffen und Unternehmensprozesse unterstützen. Unsere Ergebnisse zeigen, dass die entsprechenden Technologien inzwischen reif genug für einen Einsatz im Unternehmensumfeld sind.

Für eine noch schnellere Verbreitung im Unternehmensumfeld müssen die zugrunde liegenden Technologien jedoch zugänglicher für Nicht-Techniker werden. So sehen wir unter anderem noch Verbesserungspotenzial im Bereich der Graph-Visualisierung, des Authorings von Ontologien und Regeln und der einfachen Anbindung weiterer Datenquellen. Aufgrund der hohe Dynamik sowohl in der akademischen als auch in der angewandten Forschung (in der Siemens auch selbst aktiv ist) scheint dies jedoch eher eine Frage der Zeit. Wir sind überzeugt, dass Industrial Knowledge Graphs auch in Zukunft ein zentraler Baustein der Siemens Digitalisierungsstrategie sein werden.



Dr. Thomas Hubauer

hat in München Informatik studiert und an der Universität zu Lübeck zum Dr. rer. nat. promoviert. Seine



Dr. Steffen Lamparter

forscht seit 2009 bei Siemens Corporate Technology zur Anwendung von Datenanalyse im industriellen Umfeld und

Forschungsschwerpunkte liegen in den Bereichen Datenanalyse, Semantische Technologien und Knowledge Graphs. Innerhalb von Siemens Corporate Technology verantwortet er das Portfolio "Knowledge Graphs and Semantics".

E-Mail: thomas.hubauer(at)siemens.com

hat zahlreiche interne sowie Förderprojekte in diesem Bereich verantwortet. Seit 2016 leitet er die Forschungsgruppe "Semantics and Reasoning" im Technologiefeld "Business Analytics and Monitoring". Davor war er als Projektleiter am Institut AIFB der Universität Karlsruhe und als Geschäftsführer am Karlsruhe Service Research Institute (KSRI) tätig.

E-Mail: steffen.lamparter(at)siemens.com



Dr. Peter Haase

ist seit zwei Jahrzehnten in der Forschung und Entwicklung von semantischen Technologien an der Schnittstelle von Künstlicher Intelligenz, Wissensrepräsentation und Datenmanagement aktiv. 2006 promovierte er auf diesem Gebiet am Karlsruher Institut für Technologie (KIT). Im Jahr 2014 gründete er die metaphacts GmbH, ein Softwareunternehmen, welches die metaphactory Knowledge Graph platform entwickelt.

E-Mail: ph(at)metaphacts.com



Dr.-Ing. Daniel Herzig-Sommer

leitet als COO der metaphacts GmbH die Einführung von Knowledge Graph-Technologien in mehreren Konzernen. Er hat langjährige Erfahrung in der Softwareentwicklung und Forschung in mehreren IT-Unternehmen und Forschungsinstituten gesammelt. Er promovierte im Jahr 2013 am Karlsruher Institut für Technologie (KIT) im Bereich Semantic Search und entwickelte die Graphensuchmaschine GraphScope, die heute Teil der metaphactory Knowledge Graph platform ist.

E-Mail: dh(at)metaphacts.com

Bildnachweise:

Siemens Corporate Technology