IoTCrawler: Eine Suchmaschine für das Internet der Dinge

Marten Fischer

Hochschule Osnabrück

m.fischer@hs-osnabrueck.de

17.05.2018

Übersicht

Motivation

- Steigender Vernetzungsgrad von Geräten zum Internet der Dinge
- 50 Milliarden IoT Geräte im Jahr 2020
- Datenaufkommen übersteigt 1 Zettabyte [ABI Research]

- IoT Ressourcen müssen effizient gefunden, durchsucht und darauf zugegriffen werden können
- Großer Anteil der Entwicklungszeit benötigt für die Integration [Gartner, 2017]
- Adaptive, dynamische Lösungen zur Integration verteilter IoT Ressourcen & Daten erforderlich

Hemmnisse in dem Internet der Dinge

- Heterogene Datenquellen
- Keine Automatismen zum Finden (neuer) IoT Ressourcen
- Daten werden nicht domänenübergreifend genutzt
  - Existenz nicht bekannt ist
  - Können nicht interpretiert werden
- Keine Standards zur Verknüpfung von Rohdaten mit Metadaten
- Hohe Dynamik erschweren das Crawling, Auffinden und Bewerten von IoT Ressourcen/Daten
- Komplexe Zugriffsmechanismen behindern die Entwicklung neuer IoT Anwendungen
- Keine oder unzureichende Datenschutzkonzepte

Hemmnisse in dem Internet der Dinge

- Heterogene Datenquellen
- Keine Automatismen zum Finden (neuer) IoT Ressourcen
- Daten werden nicht domänenübergreifend genutzt
  - Existenz nicht bekannt ist
  - Können nicht interpretiert werden
- Keine Standards zur Verknüpfung von Rohdaten mit Metadaten
- Hohe Dynamik erschweren das Crawling, Auffinden und Bewerten von IoT Ressourcen/Daten
- Komplexe Zugriffsmechanismen behindern die Entwicklung neuer IoT Anwendungen
- Keine oder unzureichende Datenschutzkonzepte

$\rightarrow$ Suchmaschine zum Finden & Integration IoT Ressourcen benötigt

Anforderungen an eine Suchmaschine für das Internet der Dinge

Anforderungen an eine Suchmaschine für das Internet der Dinge

- Adaptives, verteiltes Framework
  - Abstraktion heterogener IoT Ressourcen
  - Dynamische Integration von IoT Ressourcen

Anforderungen an eine Suchmaschine für das Internet der Dinge

- Adaptives, verteiltes Framework
  - Abstraktion heterogener IoT Ressourcen
  - Dynamische Integration von IoT Ressourcen
- Skalierende Methode für das
  - Suchen (Crawling)
  - Auffinden (Discovery)
  - Indizieren (Indexing)
  - Bewerten (Ranking)

Anforderungen an eine Suchmaschine für das Internet der Dinge

- Adaptives, verteiltes Framework
  - Abstraktion heterogener IoT Ressourcen
  - Dynamische Integration von IoT Ressourcen
- Skalierende Methode für das
  - Suchen (Crawling)
  - Auffinden (Discovery)
  - Indizieren (Indexing)
  - Bewerten (Ranking)
- Möglichkeiten für eine maschinell initiierte Suche
  - Berücksichtigung des Kontext der Domänen

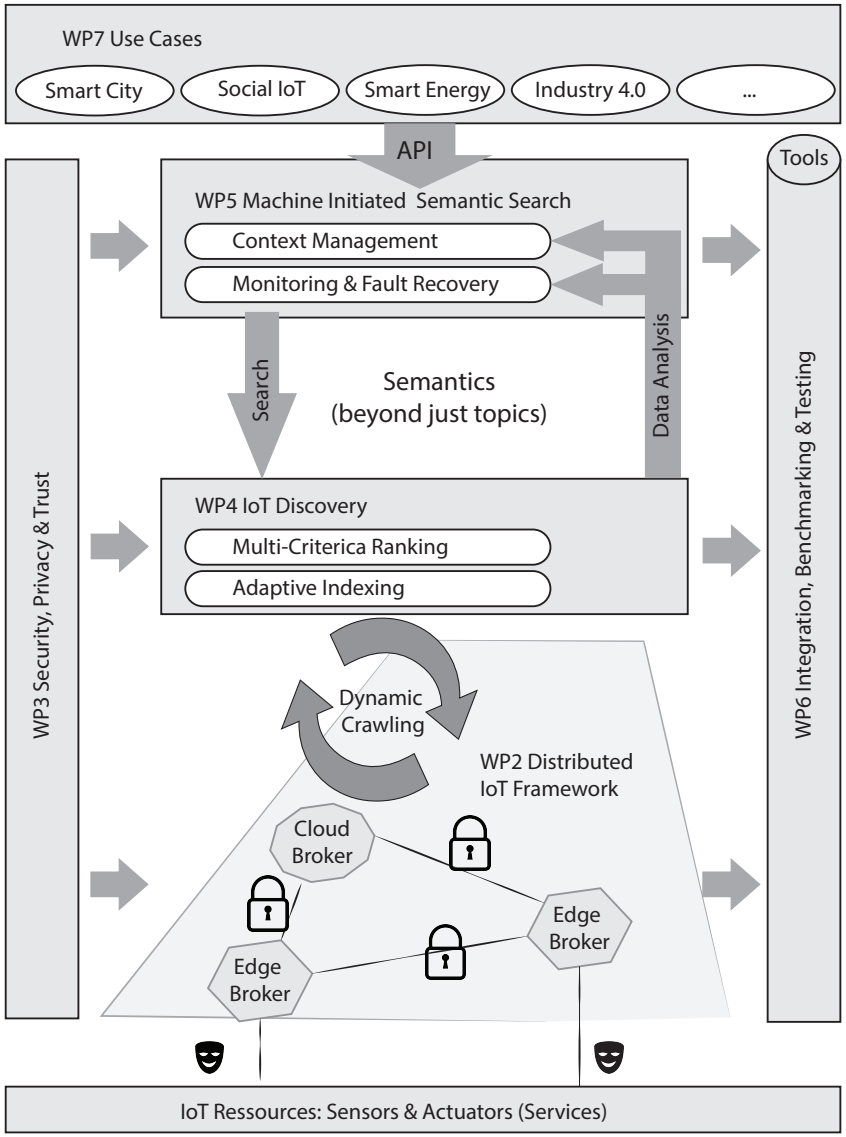
Anforderungen an eine Suchmaschine für das Internet der Dinge

- Adaptives, verteiltes Framework
  - Abstraktion heterogener IoT Ressourcen
  - Dynamische Integration von IoT Ressourcen
- Skalierende Methode für das
  - Suchen (Crawling)
  - Auffinden (Discovery)
  - Indizieren (Indexing)
  - Bewerten (Ranking)
- Möglichkeiten für eine maschinell initiierte Suche
  - Berücksichtigung des Kontext der Domänen
- Monitoring der
  - Dienstgüte (QoS)
  - Informationsqualität (QoI)
  - Input für das Bewerten

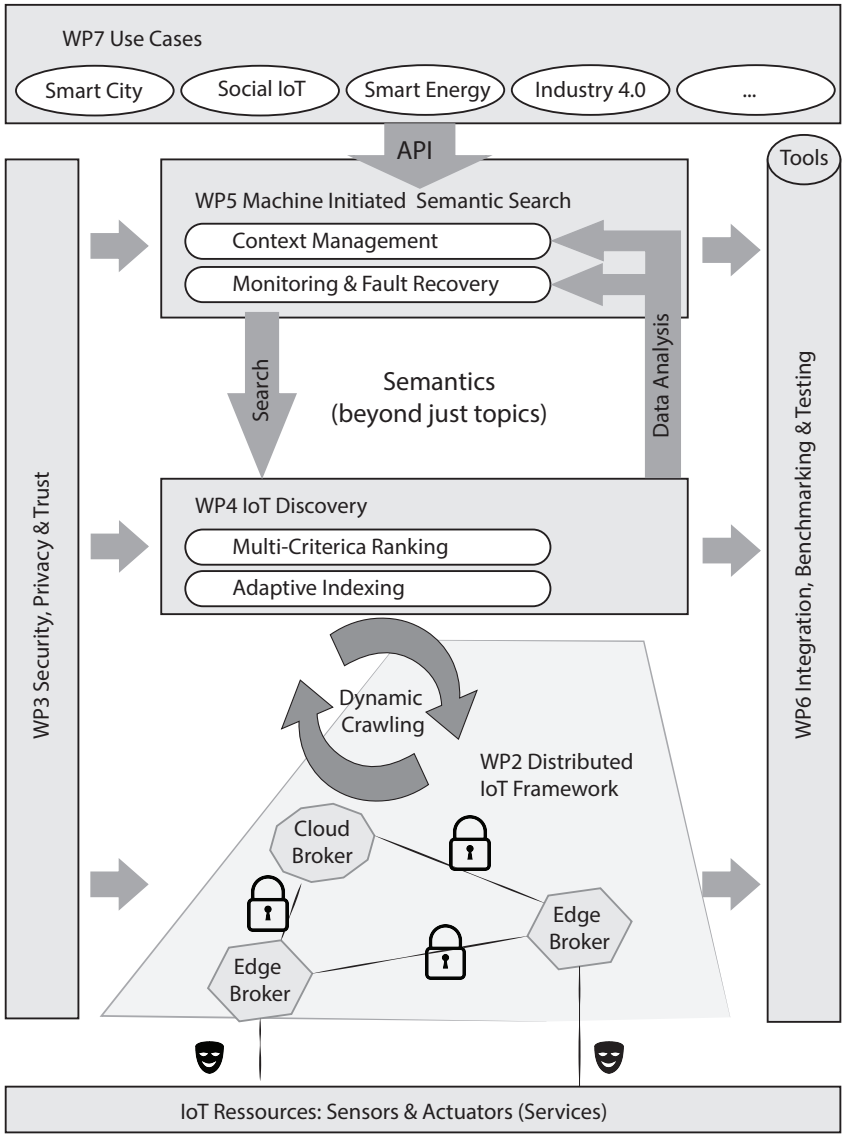
Anforderungen an eine Suchmaschine für das Internet der Dinge

- Adaptives, verteiltes Framework
  - Abstraktion heterogener IoT Ressourcen
  - Dynamische Integration von IoT Ressourcen
- Skalierende Methode für das
  - Suchen (Crawling)
  - Auffinden (Discovery)
  - Indizieren (Indexing)
  - Bewerten (Ranking)
- Möglichkeiten für eine maschinell initiierte Suche
  - Berücksichtigung des Kontext der Domänen
- Monitoring der
  - Dienstgüte (QoS)
  - Informationsqualität (QoI)
  - Input für das Bewerten
- Security by Design und Privacy by Design in allen Prozessen

IoTCrawler Systemarchitektur

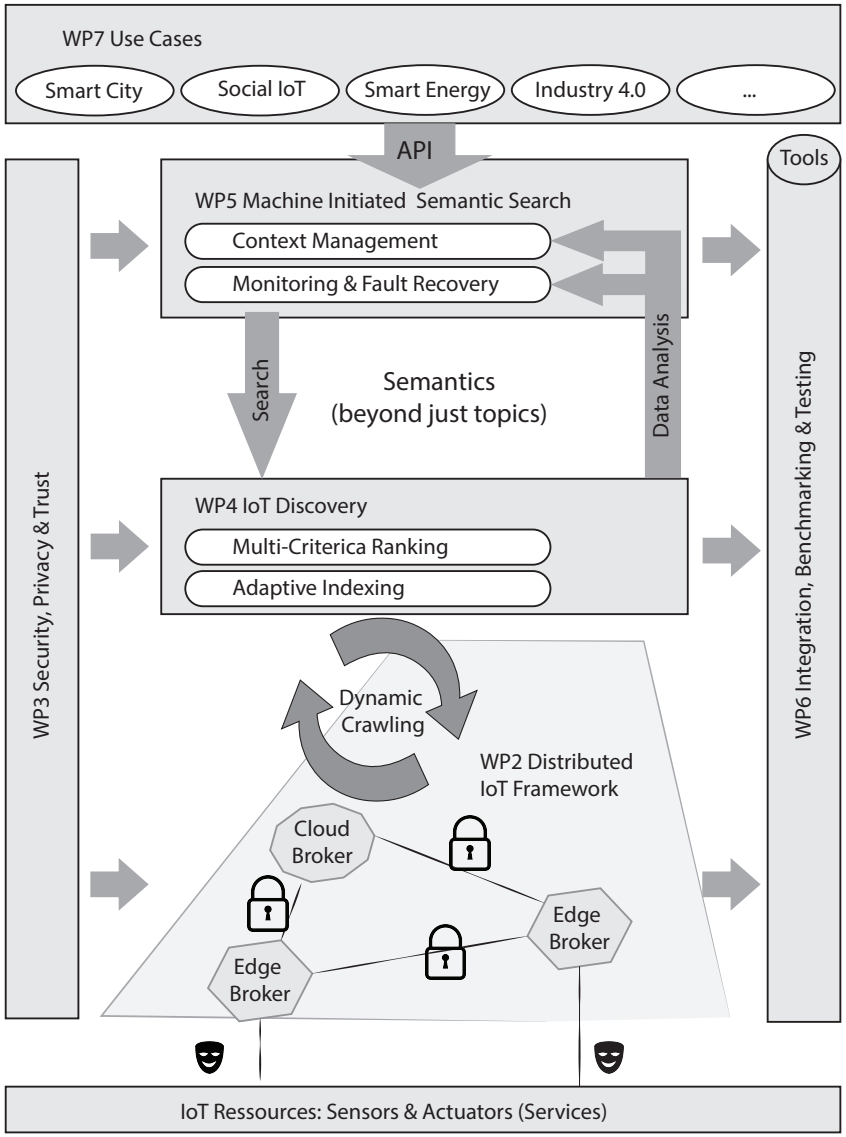


IoTCrawler Systemarchitektur



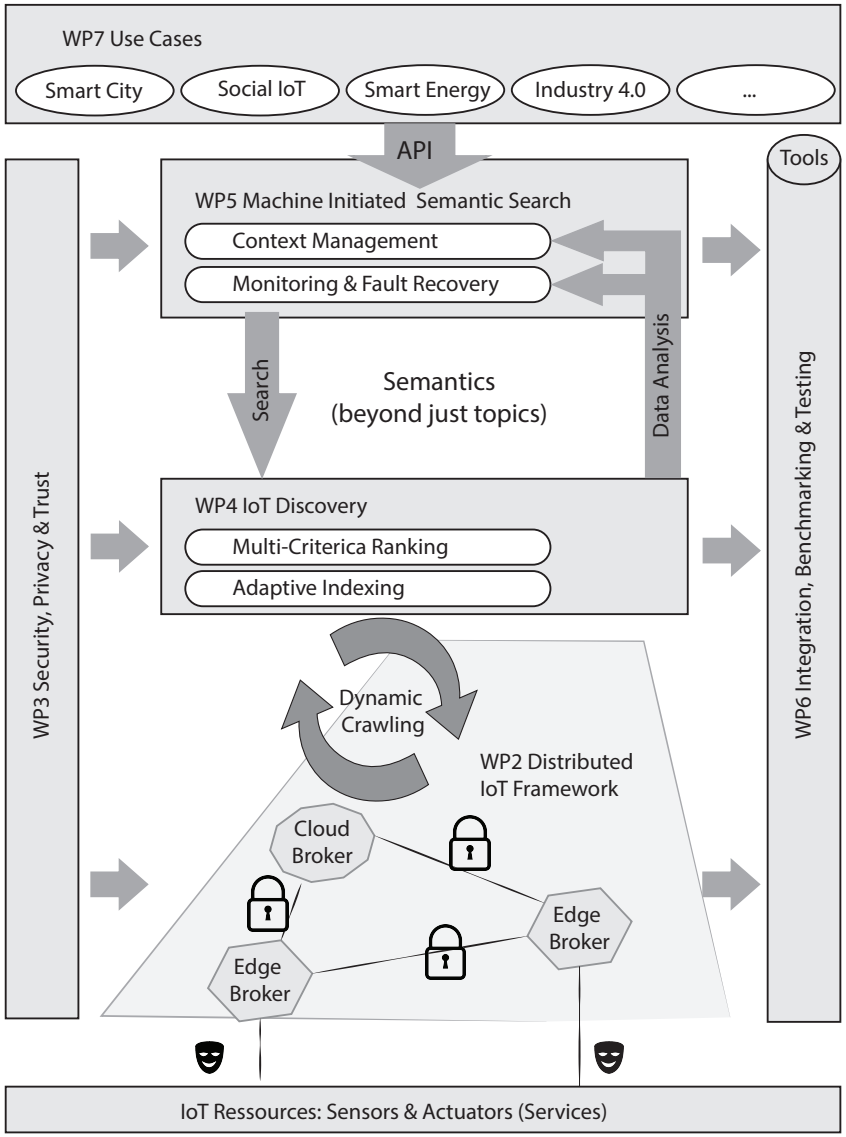
- Bereitstellen gemeinsamer Schnittstellen zur domänenübergreifenden Nutzung von IoT Ressourcen

IoTCrawler Systemarchitektur



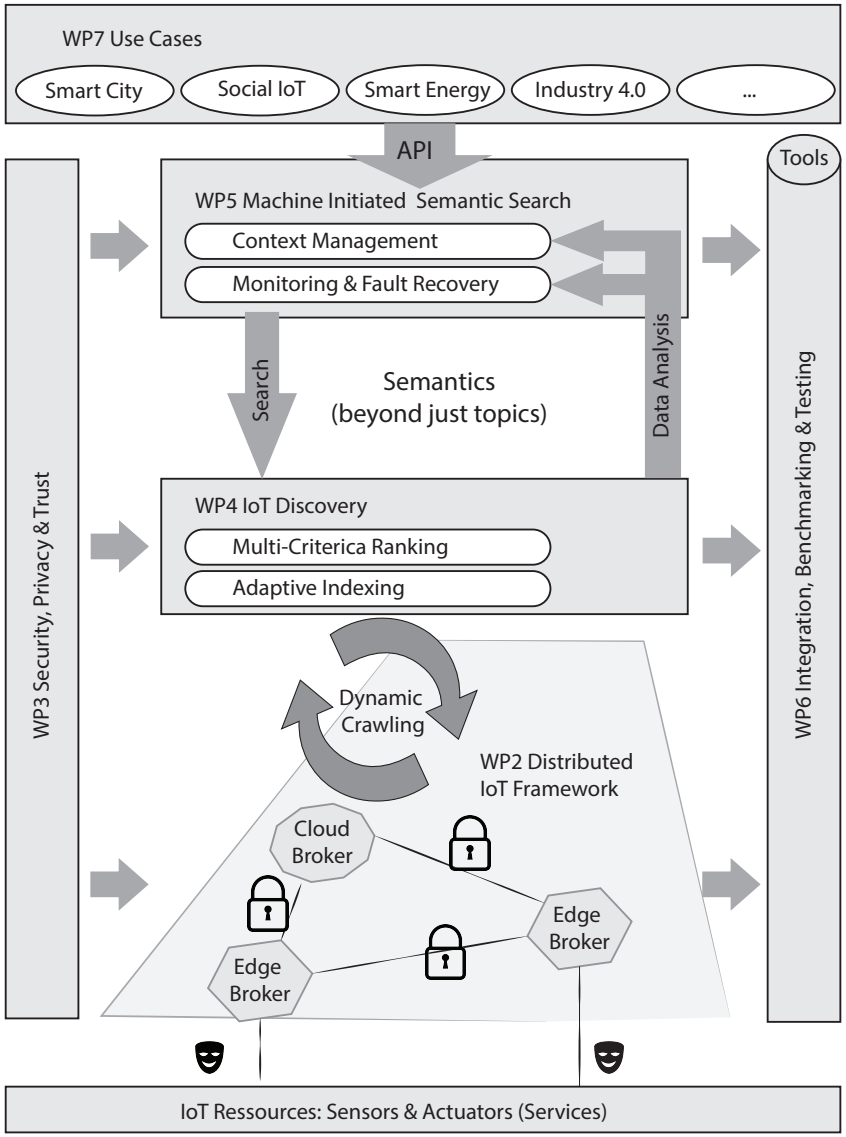
- Bereitstellen gemeinsamer Schnittstellen zur domänenübergreifenden Nutzung von IoT Ressourcen
- Skalierbare Mechanismen für das Crawling, Finden, Bewerten und Indizieren

IoTCrawler Systemarchitektur



- Bereitstellen gemeinsamer Schnittstellen zur domänenübergreifenden Nutzung von IoT Ressourcen
- Skalierbare Mechanismen für das Crawling, Finden, Bewerten und Indizieren
- Semantisch annotierte Daten zur maschinellen Suche nach Informationen/Ressourcen

IoTCrawler Systemarchitektur



- Bereitstellen gemeinsamer Schnittstellen zur domänenübergreifenden Nutzung von IoT Ressourcen
- Skalierbare Mechanismen für das Crawling, Finden, Bewerten und Indizieren
- Semantisch annotierte Daten zur maschinellen Suche nach Informationen/Ressourcen
- Zugriff auf Ressourcen unter Berücksichtigung der Datensicherheit & Schutz der Privatsphäre

Suchen und Auffinden

- Methoden zum Suchen und Finden Schlüsselkomponenten des IoTCrawler Projektes
- Distributed Hash Tables
  - Dezentralisiertes Overlay Netzwerk
  - Flexible Möglichkeiten zur Speicherung und Abfrage
  - Hohe Ausfallsicherheit
- Dynamik im Internet der Dinge erfordert Einsatz von Beschreibungsformaten
  - Resource Description Framework (RDF)
  - JSON for Linked Data (JSON-LD)
  - Abfragesprache SPARQL
- Gemanagte Verfahren für domänenübergreifenden Zugriff

Datenschutz und Datensicherheit in dem IoT

- Sicherer Informationsaustausch Attribute-Based Encryption
  - Geeignet zum Informationsaustausch mit Gruppen
  - Public-Key Verschlüsselungsverfahren
  - Schlüssel sind beschreibenden Attributen assoziiert
  - Attribute können als Zugriffsrichtlinie formuliert werden

Datenschutz und Datensicherheit in dem IoT

- Sicherer Informationsaustausch Attribute-Based Encryption
  - Geeignet zum Informationsaustausch mit Gruppen
  - Public-Key Verschlüsselungsverfahren
  - Schlüssel sind beschreibenden Attributen assoziiert
  - Attribute können als Zugriffsrichtlinie formuliert werden

- Distributed Capability-Based Access Control (DCapBAC)
  - Zugriffsrechte für ein "Smartes Objekt" mit öfftl. Schlüssel verknüpft
  - Zugriffsrechte können von Aktionen auf, beispielsweise CoAP Methoden, gemappt werden
  - $\rightarrow$ sicherer Zugriff auf Ressourcen

Datenschutz und Datensicherheit in dem IoT

- Sicherer Informationsaustausch Attribute-Based Encryption
  - Geeignet zum Informationsaustausch mit Gruppen
  - Public-Key Verschlüsselungsverfahren
  - Schlüssel sind beschreibenden Attributen assoziiert
  - Attribute können als Zugriffsrichtlinie formuliert werden

- Distributed Capability-Based Access Control (DCapBAC)
  - Zugriffsrechte für ein "Smartes Objekt" mit öfftl. Schlüssel verknüpft
  - Zugriffsrechte können von Aktionen auf, beispielsweise CoAP Methoden, gemappt werden
  - $\rightarrow$ sicherer Zugriff auf Ressourcen

- Sicheren Umgebung zur Ausführung von Transaktionen
  - Sichere/Unverfälschte Ausführung von Systemoperationen
  - Policy Monitoring & Policy Enforcement Points (PEP) kombiniert mit Blockchain-Transaktionen

Datenqualitätsanalyse

- Bewerten (Ranking) benötigt Angaben zur QoS und QoI zu den Datenquellen
- Grundsätzliches Problem Fehlen von Wissen über tatsächliche Werte (Ground-Truth)
- IoTCrawler verfolgt zwei Phasen Ansatz
  - Atomic Monitoring: sensorspezifische Eigenschaften (z.B. Einhalten der Abtastfrequenz)
  - Composite Monitoring: Validierung der Messwerte durch benachbarte Sensoren
  - Erfordert infrastrukturbasierte Distanz- und Korrelationsmodell
    - ★ hier domänenübergreifend
  - Nutzt IoTCrawler's Methoden zum Suchen und Finden von IoT Ressourcen

Domäne I: Smart City

- Umsetzung in der dänischen Stadt Aarhus
- Technologien zum Finden neuer Datenquellen für Open Data Plattform (Open Data DK)
- Bereitstellen von Werkzeugen für "City Lab" um negativer Wahrnehmung bzgl. IoT entgegenzuwirken
- Einbindung von Bürgern und Firmen $\rightarrow$ Heranführen an die Welt des IoT
- Monitoring von Aktivität und Qualität zur Überwachung der Performanz
  - KPIs für Aarhus City Lab

Domäne II: Social IoT

- Bestimmen und Verbessern von Zuschauererfahrungen auf Events durch Sensoren
- Austragungsorte mit Sensoren bestückt; Zuschauer tragen Wearables
- über 800 Events aus unterschiedlichen Bereichen verfügbar
- Generieren neuer Inhalte und Erkenntnisse in Kombination mit Social Media
- Steigerung der Nutzbarkeit durch Technologien zum Finden und Annotieren von Daten(quellen)
- $\rightarrow$ größerer Nutzerkreis zugänglich; neuartige Anwendungen möglich

Domäne III: Smart Energy

- Energiewende $\rightarrow$ zunehmend dezentrale Strukturen im Energienetz
- Haushalte nicht mehr nur Verbraucher sondern auch Erzeuger
- Volatile Erzeugung (Wind und Sonne)
- Kommunikation Haushalt $\leftrightarrow$ Smart Grid zur Stabilisierung des Netzes
  - Austausch Informationen über (variabler) Verbraucher und Erzeuger
  - Handel mit Energie

- IoTCrawler Technologien:
  - Crawling: Entdecken von Haushaltsgeräten
  - Indizieren: Analyse der Anforderungen/Potential zur Bereitstellung von Energie
  - Suche: beispielsweise Möglichkeit zur Suche nach Verbrauch (nicht Gerät) für ein Netzbetreiber

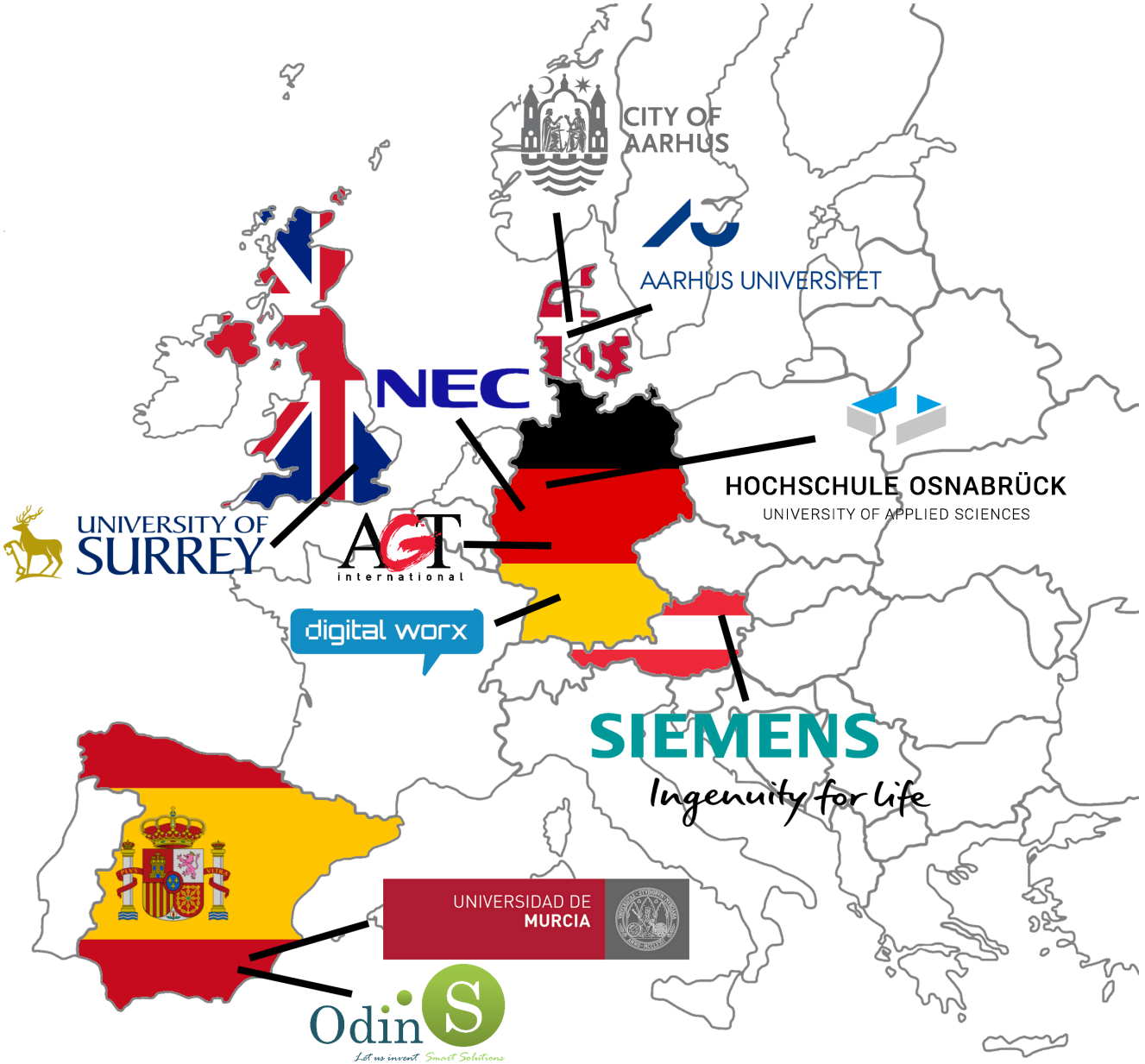
Domäne IV: Industrie 4.0

- Integration neuer Datenquellen in Industrie 4.0 Analysesoftware 80% der aufgewendeten Zeit
- Beschleunigung durch IoTCrawler Technologien zum Finden von:
  - Maschinen-Metadaten
  - Sensordaten
  - Informationen aus Unternehmensdatenbanken
- Durchgängiges Monitoring der Datenströme zur frühzeitigen Erkennung von Fehlerzuständen

Zusammenfassung - IoTCrawler Enablers

- Indizierung und Bewertung
  - Bewertung der IoT Ressourcen/Informationen anhand der QoI/QoS
  - Atomic- und Composite-Monitoring
  - Skalierbare Indizierung der Daten(-quellen) über geeignete Parameter
- Maschinelle Suche
  - Genutzt von Anwendungen, Diensten und Geräten
  - Suche in den indizierten, semantisch annotierten Daten
  - Einbeziehen des Kontextes von Ressource und Anwendung
- Domänenübergreifender Zugriff auf IoT Ressourcen
  - Gemeinsame Schnittstellen für Zugriff aus unterschiedlichen Domänen
  - Ansätze werden in 4 unterschiedlichen Domänen evaluiert
- Ganzheitlicher Datenschutz & Datensicherheit
  - Kryptographisch gesicherter Informationsaustausch mit Gruppen (ABE)
  - Gesicherte Ausführung von Systemoperationen (Blockchain)
  - Kopplung von Zugriffsrechten mit öfftl. Schlüsseln (DCapBAC)

IoTCrawler Konsortium



Fragen?

https://www.iotcrawler.eu

Referenzen I

Adi Shamir (1984)

Identity-based cryptosystems and signature schemes

Theory and Application of Cryptographic Techniques pages 47 - 53. Springer, 1984.

ABI Research, "Data Captured by IoT Connections to Top 1.6 Zettabytes in 2020"

https://www.abiresearch.com/press/data-captured-by-iot-connections-to-top-16-zettaby/

Eric Thoo, Ted Friedmann "IoT Data Proliferation Elevates Data Integration Challenges"

https://www.gartner.com/doc/3221917/iot-data-proliferation-elevates-data