Internet of Things and Data Science

Chapter · February 2022

DOI: 10.5281/zenodo.5901124

Petar Kocovic
Univerzitet UNION Nikola Tesla

Chapter 1
When Data Science Meets the Internet of Things

Petar Kočović
University Union "Nikola Tesla" – School for Information Technology and Engineering – Belgrade, Serbia

Božidar Radenković
University of Belgrade, School for Organizational Sciences, Belgrade, Serbia

DOI 10.5281/zenodo.5901124

ABSTRACT

Rapid developments in hardware, software, and communication technologies, software tools, and algorithms for big data processing have facilitated the emergence of Internet-connected sensory devices that provide observations and data measurements from the physical world. By 2025, it is estimated that the total number of Internet-connected devices being used will be between 50 and 75 billion. As these numbers grow and technologies become more mature, the volume of data being published will increase. The technology of Internet-connected devices, referred to as the Internet of Things (IoT), continues to extend the current Internet by providing connectivity and interactions between the physical and cyber worlds. In addition to an increased volume, the IoT generates big data characterized by its velocity in terms of time and location dependency, with a variety of multiple modalities and varying data quality. Intelligent processing and analysis of these big data are the keys to developing smart IoT applications. This article assesses the various machine learning methods that deal with the challenges presented by IoT data by considering smart cities as the main use case. The key contribution of this study is the presentation of a taxonomy of machine learning algorithms explaining how different techniques are applied to the data to extract higher-level information. The potential and challenges of machine learning for IoT data analytics will also be discussed. A use case of applying a Support Vector Machine (SVM) to Aarhus smart city traffic data is presented for a more detailed exploration.

Keywords: applied machine learning and modeling, ARIMA, big data and data engineering, business intelligence, data analytics, data monetization, data science education, data science in the Fourth industrial revolution, deep learning and artificial intelligence, digital transformation, open data, Bagging Classification, Canonical Correlation Analysis, Classification, Regression Trees Classification, Density-Based Spatial, Clustering of Applications with Feed Forward Neural Network, K-means, K-Nearest Neighbors, Linear Regression, Naive Bayes, One-class Support Vector Machines, Principal Component Analysis, Random Forests Classification, Support Vector Machine, Support Vector Regression

INTRODUCTION

The Internet of Things (IoT) is enabled by heterogeneous technologies used to sense, collect, store, act, process, infer, transmit, create notifications, manage and analyze data. The combination of emergent technologies for information processing and distributed security, e.g. AI, IoT, DLT's and blockchain, brings new challenges in addressing distributed IoT architectures and distributed security mechanisms that form the foundation of improved and, eventually entirely new products and services.

IoT developments in previous years have been characterized by attributes that can be "labeled" wit 6As: **Anything** (Any device), to be transferred from/to **Anyone** (anybody), located **Any place** (anywhere) at **Any time**, using **Any** (physical) **path** (any network) to provide **Any service** (suitable for any business). (I.I., 2005) (Sade Kuyoro, Folasade Osisanwo, Omoyele Akinsowon, 2015)

Figure 1. IoT as 6As

IoT ecosystems are built on 6Cs elements: Collect (from devices with various complexities, heterogeneous), Connect, Cache, Compute, Cognize (based on information analytics, insights, extractions, real-time AI processing), and create (create of new interactions, business models and services).

The third wave of the Internet Internet of Things. There are different assumptions about the number of connected devices worldwide, and different trends of growth. Figure 3 shows the forecast form of 2015-2025. (Statista, Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025, 2020) In 2020 we expect around 31 billion connected devices; in 2025-75 billion – 2.5 times the number of IoT devices will grow in only 6 years.

Figure 2. Next-Generation IoT Hyperconnected: 6As and 6Cs

The significant rising of the IoT came because of the significant technology changes, that include the following:

Cheap sensors – average prices dropped down two times in the past 10 years (from $1.30 to 60 cents).

Cheap bandwidth – average prices dropped down forty times in the past 10 years.

Cheap processing – processing costs dropped down nearly 60 times in the past 10 years, introducing smartphones on the scene generating or receiving all the new data.

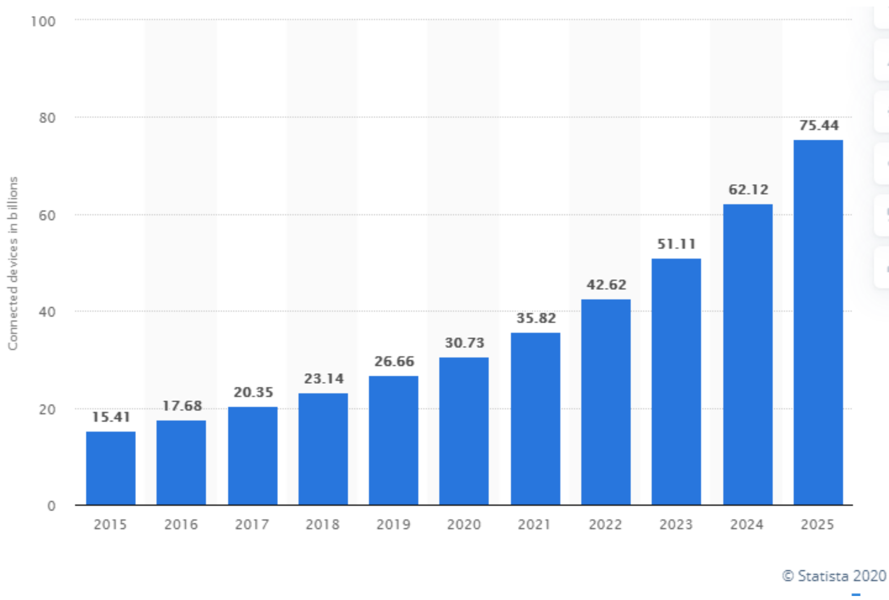

Figure 3 – IoT connected devices: Source: Statista.com

Smartphones – Smartphones are now becoming the personal gateway to the IoT, serving as a remote control or hub for the connected home, connected car, or the health and fitness devices consumers are increasingly starting to wear.
5G and EDGE technology – More than 100 times faster Internet via mobile devices, can help to transfer the data immediately. New development in the EDGE technology can add more secure and fastest transfer of 5G data.
Ubiquitous wireless coverage – With Wi-Fi coverage now ubiquitous, wireless connectivity is available for free or at a very low cost, given Wi-Fi utilizes an unlicensed spectrum and thus does not require monthly access fees to a carrier.
Big data – As the IoT will by definition generate voluminous amounts of unstructured data, the availability of big data analytics is a key enabler.

IPv6 – Most networking equipment now supports IPv6, the newest version of the Internet Protocol (IP) standard that is intended to replace IPv4. IPv4 supports 32-bit addresses, which translates to about 4.3 billion addresses – a number that has become largely exhausted by all the connected devices globally. In contrast, IPv6 can support 128-bit addresses, translating to approximately 3.4 x 1038 addresses – an almost limitless number that can amply handle all conceivable IoT devices. At the beginning of November 2019, the IPv4 protocol stops to address new devices. For compatibility purposes, devices addressed with IPv4 can be used parallel with IPv6 in the future.
Digitalization – Digitalization is being utilized in many fields and industrial sectors.

Concepts like Smart home and Smart City now come to life. Today we have a lot of household gadgets such are smart refrigerators, washing machines, TV sets, boiler thermostats, door openers. This fast development was possible because of forming IoT Alliances and associations. Devices use few industrial standards.

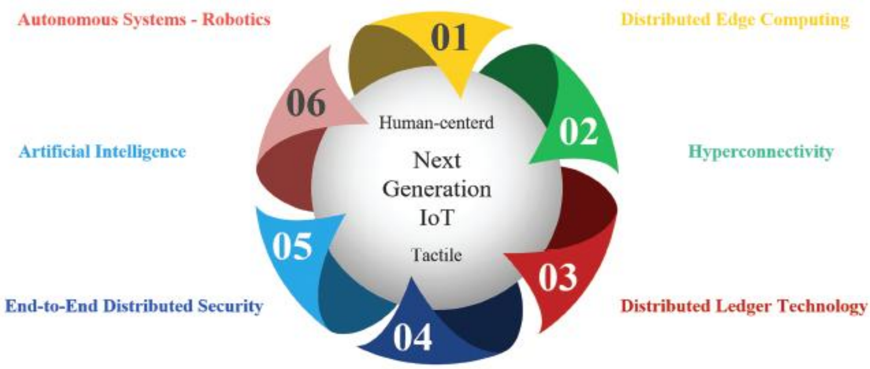

Figure 4 - Next generation of IoT

This leads to the development of the Next generation of IoT, figure 4. International Telecommunication Union defined the Internet of Things in the Global Standard Initiative. (ITU, 2020)

IOT STANDARDS, DATA LINKS, PROTOCOLS, AND ALLIANCES

Authors found inspiration for this section in the paper "A Survey of Protocols and Standards for the Internet of Things". (Tara Salman, Raj Jain, 2017). But we restructure standards and protocols in different ways. Instead of trying to organize the IoT Protocol like ISO/OSI Model, we have broken the protocols into the following layers to provide some level of organization:

- Infrastructure (ex: 6LowPAN, IPv4/IPv6, RPL)
- Identification (ex: EPC, uCode, IPv6, URIs)
- Comms / Transport (ex: Wifi, Bluetooth, LPWAN)
- Discovery (ex: Physical Web, mDNS, DNS-SD)
- Data Protocols (ex: MQTT, CoAP, AMQP, Websocket, Node)
- Device Management (ex: TR-069, OMA-DM)
- Semantic (ex: JSON-LD, Web Thing Model)
- Multi-layer Frameworks (ex: Alljoyn, IoTivity, Weave, Homekit)

Security

Industry Vertical (Connected Home, Industrial, etc)

Infrastructure

IPv6 - IPv6, is an Internet Layer protocol for packet-switched internetworking and provides end-to-end datagram transmission across multiple IP networks. IPv6 protocol was produced by the Internet Engineering Task Force (IETF). (IPv6, 2020)

6LoWPAN is an acronym for IPv6 over Low power Wireless Personal Area Networks. It is an adaption layer for IPv6 over IEEE802.15.4. This protocol operates only in the 2.4 GHz frequency range with a 250kbps transfer rate. (IPv6 over Low power WPAN (6lowpan), 2020)

UDP (User Datagram Protocol) - UDP is the main alternative to TCP and one of the oldest network protocols in existence, introduced in 1980. UDP is often used in applications specially tuned for real-time performance. (Postel, 2020)

QUIC (Quick UDP Internet Connections, pronounced quick) supports a set of multiplexed connections between two endpoints over User Datagram

Protocol (UDP) and was designed to provide security protection equivalent to TLS/SSL. (J. Iyengar, M. Thomson, 2020)

uIP - open-source TCP/IP used with tiny 8 and 16-bit microcontrollers, developed by Adam Dunkels (uIP (Micro IP), 2020)

DTLS (Datagram Transport Layer) - provides communications privacy for datagram protocols. DTLS protocol is based on Transport Layer Security (TLS). (E. Rescorla, N. Modadugu, 2020)

ROLL – IPv6 routing for low power networks (Routing Over Low power and Lossy networks (roll), 2020)

nanoIP - is a concept that was created to bring Internet-like networking services to embedded and sensor devices, without the overhead of TCP/IP, and with minimal overheads, wireless networking, and local addressing. (About nanoIP, 2020)

Content-Centric Networking (CCN) - Next-gen network architecture to solve challenges in content distribution scalability, mobility, and security. (Muscanello, 2020)

Time Synchronized Mesh Protocol (TSMP) - communications protocol for self-organizing networks of wireless devices called motes.

Discovery

- **mDNS (multicast Domain name System)** - provides the ability to perform DNS-like operations on the local link in the absence of any conventional Unicast DNS server. (S. Cheshire, M. Krochmal, 2020)
- **Physical Web** – enables everyone to see a list of URLs with a Bluetooth Low Energy (BLE) beacon.
- **BSI group** – British Standard institution that provides valuable standards and education in the Internet of Things. (Internet of Things, 2020)
- UPnP (Universal Plug and Play) – From January 1, 2016, is part of the Open Connectivity Foundation. UPnP technology allows devices to connect seamlessly and to simplify network implementation in the home and corporate environments. An industry initiative that gained more than 1000 leading companies in computing, printing, and networking; consumer electronics; home appliances, automation, control and security; and mobile products. (About UPnP, 2020)

Data Protocols

AMQP (Advanced Message Queuing Protocol) – AMQP 1.0 approved as an international standard. (AMQP, 2020)

CoAPP (The Constrained Application Protocol). The Constrained Application Protocol (CoAP) is a specialized web transfer protocol for use with constrained nodes and constrained (e.g., low-power, lossy) networks. (The Constrained Application Protocol (CoAP), 2020)

DDIS (Data-Distribution Service for Real-Time Systems) – DDIS is a mature proven open International standard for Internet connectivity, ideal for the Internet of Things. (2020)

HTTP/2 – HTTP version 2 (M. Belshe, R. Peon, M. Thomson, 2020)

JMS – Java Message Service

LLAP (Lightweight Local Automation Protocol) (LLAP protocol spec #6, 2020)

LWM2M (Lightweight M2M) - Lightweight M2M (LWM2M) is a system standard in the Open Mobile Alliance (LWM2M – M2M The Lighter Way, 2020)

Mihini/M3DA – This agent act as a software component that acts as a mediator between an M2M server and the applications running on an embedded gateway (Mihini/M3DA Specification, 2020)

MQTT-SN (MQTT for Sensor Networks) - MQTT is a machine-to-machine (M2M)/"Internet of Things" connectivity protocol. It was designed as an extremely lightweight publish/subscribe messaging transport. (MQTT, 2020)

Mosquitto –an open-source MQTT protocol, versions 5.0, 3.1.1, and 3.1 (Eclipse Mosquitto, 2020)

ONS (Object Name Service) – Standard (Object Name Service (ONS), 2020)

Reactive Streams - Reactive Streams is an initiative to provide a standard for asynchronous stream processing with non-blocking backpressure. (Reactive Streams, 2020)

REST (Representational State Transfer) - Software architectural style that defines a set of constraints to be used for creating Web services. (Representational state transfer, 2020)

SMCP – A C/based CoAPP stack. (Gitub, 2020)

SOAP (Simple Object Access protocol) – W3 protocol (Latest SOAP versions, 2020)

SSI (Simple Sensor Interface Protocol) - a simple communications protocol designed for data transfer between computers or user terminals and smart sensors (Simple Sensor Interface protocol, 2020)

STOMP – The Simple Text Oriented Messaging Protocol (STOMP Servers, 2020)

XMPP (Extensible Messaging and Presence Protocol) - an open standard for messaging and presence. (The universal messaging standard, 2020)

XMPP-IoT – XMPP standard for Internet of Things (XMPP-IoT, 2020)

WebSocket - The WebSocket specification—developed as part of the HTML5 initiative. (Websocket, 2020)

Communication/Transport Layer

Table 1 shows available technologies based on the following criteria: Development Environment: a) Data rate, b) Range (network coverage), c) Security, d) Power consumption, e) Network topology, f) Network directionality (uni-directional vs bi-directional), g) Network size, h) Network reliability, i) Backhaul, j) Available spectrum, k) Standards licensing, l) Available parts and m) Cost.

Table 1. Categories vs available technologies. Modified and extended from (Vidales, 2020)

Technology	Frequency	Date Rate	Range	Power usage	Cost
5G	Microwave cellular (1-6GHz)	10Gb/sec	Short	High	High
4G	Radio cellular (700-2,700MHz)	300Mb/sec	Several Kilometers	High	High
3G	Radio cellular	10Mb/sec	Several Kilometers	High	High
Bluetooth/BLE	2.4GHz	1/3 Mb/s	~ 1 km	Low	Low
IEEE 802.15.4	Sub GHz-2GHz	/sec	>250m^2	Low	Low
LoRa	Sub GHz	< 50kb/sec	1.5-5 km	Low	Medium
LTE-Advanced	3.9 GHz	1Gb/s	Several kilometers	Medium	High
NB-IoT	Cellular bands	0.1-1 Mb/Sec	Several kilometers	Medium	High
WiFi	SubGHZ-2.4GHz, 5GHz	< 54MB/sec	< 100m	Medium	Low
ZigBee	2.4GHz	250kb/sec	~100m	Low	Medium
Z-Wave	SubGHZ	40kb/sec	~100m	Low	Medium

ANT (Adaptive Network Topology) - multicast wireless sensor network technology (ANT, 2020)

Bluetooth - works in the 2.4 GHz ISM band uses frequency hopping, with a data rate of up to 3 Mbps and a maximum range of 100m. (Bluetooth, 2020)

DigiMesh - DigiMesh is a proprietary wireless mesh networking topology developed by Digi's RF engineering experts which allows for time-synchronized sleeping nodes and low-power operation. (Products Tagged "DigiMesh", 2020)

EC-GSM-IoT (Extended Coverage – GSM – IoT) - (formerly EC-EGPRS) is a new technology that enables new capabilities of existing cellular networks for LPWA (Low Power Wide Area) IoT applications. (Ericsson and Orange in the Internet of Things trial with EC-GSM-IoT, 2020)

Edystone - A protocol that defines a Bluetooth low energy (BLE) message format for proximity beacon messages. (Edystone, 2020)

EnOcean - technology that works in the frequencies of 868 MHz for Europe and 315 MHz for North America. (Self-powered Wireless Standard , 2020)

IEEE 802.15.4 - Defines the operation of low-rate wireless personal area networks (LR-WPANs) (IEEE 802.15 WPAN™ Task Group 4 (TG4), 2020)

ISA 100.11a - ISA100.11a is a wireless networking technology standard developed by the International Society of Automation (ISA). (Vavra, 2020)

LoRA WAN - the specification is a Low Power, Wide Area (LPWA) networking protocol designed to wirelessly connect battery operated 'things' to the internet in regional, national, or global networks, and targets key Internet of Things (IoT) requirements such as bi-directional communication, end-to-end security, mobility and localization services. (What is the LoRaWAN® Specification?, 2020)

LTE-MTC (LTE-Machine Type Communication) - Standards-based family of technologies supports several technology categories (Long Term Evolution for Machines: LTE-M, 2020)

NB-IoT (Narrowband IoT) - A Low Power Wide Area Network technology being standardized by the 3GPP (Narrowband – Internet of Things (NB-IoT), 2020)

NFC (Near Field Communication) - Based on the standard ISO/IEC 18092:2004, using inductively coupled devices. The center frequency is 13.56 MHz, the data rate is up to 424 kbps, and the range with a few meters. (NFC Forum, 2020)

WiMax - is based on the standard IEEE 802.16. The range goes up to 50 km and for mobile devices with 5 to 15 km. WiMax operates at frequencies 2.5-5.8 GHz with a transfer rate of 40 Mbps. (WiMax Forum, 2020)

ZigBee - protocol uses the IEEE 802.15.4 standard, operates in the 2.4 GHz frequency with 250 kbps. The maximum number of nodes in the network is 1024 with a range of up to 200 meters. (Control Your World, 2020)

Figure 5 represents different technologies and standards depending on the distance, data rate, and power consumption.

SEMANTIC

IOTDB is a JSON-linked data standard for describing IoT. (IOTDB, 2020)

LsDL (Lemonbeat smart Device Language) - XML-based device language for service-oriented devices (LEMONBEAT OS, 2020)

RAML (RESTful API Modeling Language) - Makes it easy to manage the whole API lifecycle from design to sharing. It's concise - you only write what you need to define - and reusable. (RAML, 2020)

Semantic Sensor Net Ontology – W3C - This ontology describes sensors and observations, and related concepts. It does not describe domain concepts, time, locations, etc. these are intended to be included from other ontologies via OWL imports. (Semantic Sensor Net Ontology, 2020)

SENML (Media Types for Sensor Markup Language) - A simple sensor, such as a temperature sensor, could use this media type in protocols such as HTTP or CoAP to transport the measurements of the sensor or to be configured. (C. Jennings, Z. Shelby, J. Arkko, 2020)

SensorML - The primary focus of the Sensor Model Language (SensorML) is to provide a robust and semantically-tied means of defining processes and processing components associated with the measurement and post-measurement transformation of observations. (Sensor Model Language (SensorML), 2020)

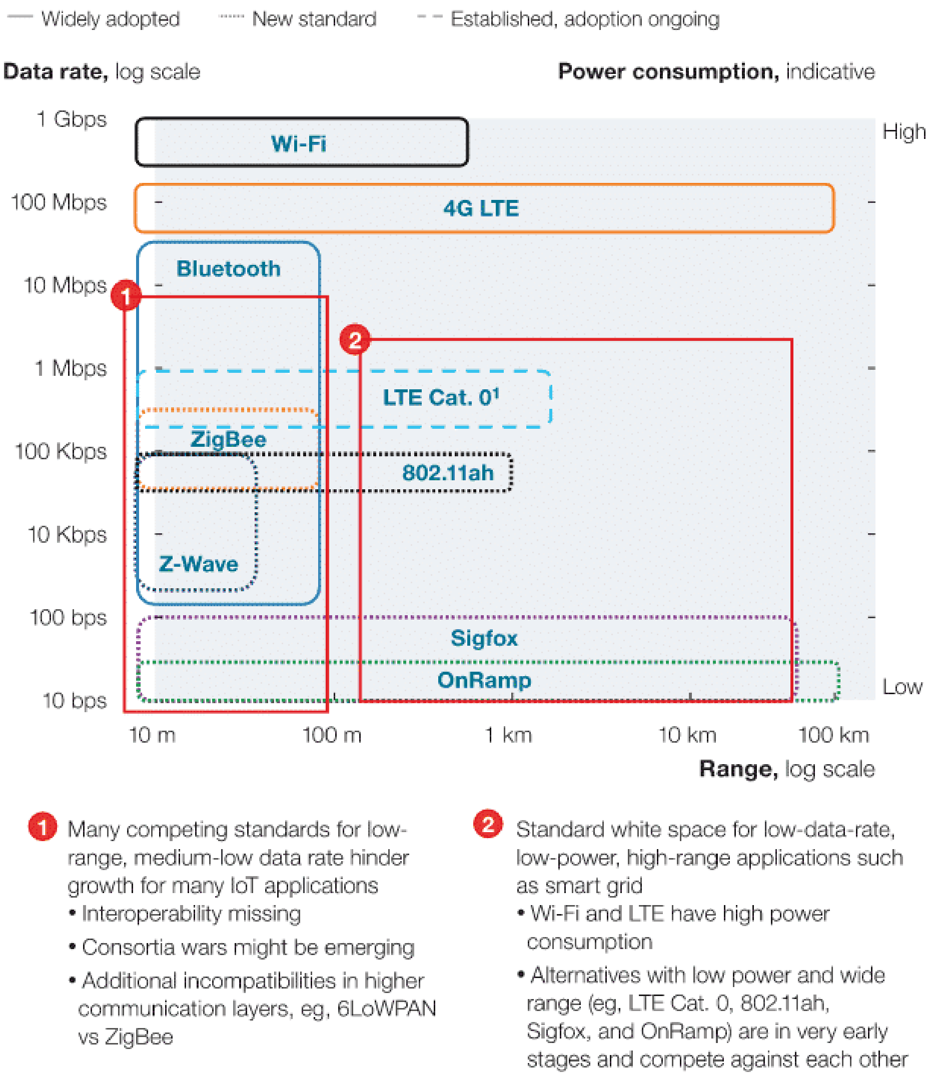

Figure 5 - Different technologies and standards depending on the distance, data rate, and power consumption.

Multilayer Frameworks

Alljoyn - An open-source software framework that makes it easy for devices and apps to discover and communicate with each other. In

October 2016 AllJoyn merged with the Open Connectivity Foundation (OCF). (OCF SOLVING THE IOT STANDARDS GAP, 2020)

IEEE P2413:2019 - This standard defines an architecture framework description for the Internet of Things (IoT) that conforms to the international standard ISO/IEC/IEEE 42010:2011. (2413-2019 - IEEE Approved Draft Standard for an Architectural Framework for the Internet of Things (IoT), 2020)

IoTivity is an open-source software framework enabling seamless device-to-device connectivity to address the emerging needs of the Internet of Things. (IoTivity, 2020)

Security

Open Trust Protocol (OTrP) is used in environments where security services should be isolated from a regular operating system (often called rich OS). (M. Pei, A. Atyeo, N. Cook, M. Yoo, H. Tschofenig, 2020)

X.509 - In cryptography, X.509 is a standard defining the format of public-key certificates. (X.509, 2020)

Vertical Specific

IEEE 1451 – Standard for Smart Transducers (John Mark, Paul Hufnagel, 2020)

IEEE 1888.3:2013 - Standard for Ubiquitous Green Community Control Network: Security (1888.3-2013 - IEEE Standard for Ubiquitous Green Community Control Network: Security, 2020)

IEEE 1905.1-2013 - Standard for a Convergent Digital Home Network for Heterogeneous Technologies (IEEE 1905.1-2013 - IEEE Standard for a Convergent Digital Home Network for Heterogeneous Technologies, 2020)

IEEE 802.16p-2012 - Standard for Air Interface for Broadband Wireless Access Systems--Amendment 1: Enhancements to Support Machine-to-Machine Applications (IEEE 802.16p-2012 - Standard for Air Interface for

Broadband Wireless Access Systems--Amendment 1: Enhancements to Support Machine-to-Machine Applications, 2020)

IEEE 1377:2012 - Standard for Utility Industry Metering Communication Protocol Application Layer (End Device Data Tables) (IEEE 1377:2012 - Standard for Utility Industry Metering Communication Protocol Application Layer (End Device Data Tables), 2020)

Alliances and Organizations

Allseen Alliance merges with Open Connectivity Foundation to Accelerate the Internet of Things (OCF, 2020)

Eclipse Paho Project - The scope of the Paho project is to provide open-source implementations of open and standard messaging protocols that support current and emerging requirements of M2M integration with Web and Enterprise middleware and applications. It will include client implementations for use on embedded platforms along with corresponding server support as determined by the community. (Paho, 2020)

EPC Global - s a GS1 initiative to innovate and develop industry-driven standards for the Electronic Product Code™ (EPC) to support the use of Radio Frequency Identification (RFID) and allow global visibility of items (EPCIS) in today's fast-moving, information-rich, trading networks. (EPCglobal, 2020)

ETSI (European Telecommunications Standards Institute) (ETSI, 2020)

HIPRG (Host Identity Protocol ResearchGroup) (Host Identity Protocol (HIP), 2020)

IETF (Internet Engineering Task Force)
CoRE working group (Constrained RESTful Environments) (Carsten Bormann, Jaime Jimenez, Alexey Melnikov, 2020)

6lowpan working group (IPv6 over Low power WPAN) (Carsten Bormann, Geoffrey Mulligan, Ted Lemon, 2020)

ROLL working group (Routing Over Low power and Lossy networks) (Routing Over Low power and Lossy networks (roll), 2020)

IEEE (Institute of Electric and Electronic Engineers) - Internet of Things (Internet of Things, 2020)

Industrial Internet Consortium - The Industrial Internet Consortium was founded in March 2014 to bring together the organizations and technologies necessary to accelerate the growth of the industrial internet by identifying, assembling, testing, and promoting best practices. (IIC, 2020)

IoT-GSI - Global Standards Initiative on the Internet of Things (Internet of Things Global Standards Initiative, 2020)

IPSO is a global non-profit organization serving the various communities seeking to establish the Internet Protocol as the network for the connection of Smart Objects by providing coordinated marketing efforts available to the general public. (Oma SpecWorks, 2020)

ISA - International Society of Automation (ISA, 2020)

ISO – International Standard Organization (ISO, 2020)

OMG (Object Management Group) (OMG, 2020)
Data Distribution Service Portal (DDS Standard, 2020)

OASIS (Organization for the Advancement of Structured Information Standards) – MQTT Technical Committee (OASIS Message Queuing Telemetry Transport (MQTT) TC, 2020)

OGC (Open Geospatial Consortium) (Sensor Model Language (SensorML), 2020)

OneM2M - The purpose and goal of oneM2M are to develop technical specifications that address the need for a common M2M Service Layer that can be readily embedded within various hardware and software, and relied upon to connect the myriad of devices in the field with M2M application servers worldwide. (Standards for M2M and the Internet of Things, 2020)

OpenWSN - Serves as a repository for open-source implementations of protocol stacks based on Internet of Things standards, using a variety of hardware and software platforms. (OpenWSN, 2020)

OSIOT - An organization with a single focus to develop and promote royalty-free, open-source standards for the emerging Internet of Things. (OSIOT, 2020)

W3C - Semantic Sensor Net Ontology (Semantic Sensor Net Ontology, 2020)
Web of Tings Comunity Group (WEB OF THINGS COMMUNITY GROUP, 2020)

DIGITAL DISRUPTION CHANGING ENTERPRISE ENVIRONMENT

The Internet of Things (IoT) is emerging as the third wave in the development of the Internet. The first Internet wave connected 1 billion users in 2005, while the 2012 mobile wave connected another billion. Figure 6. (Statista, Half of the World will be on the World Wide Web, 2020)

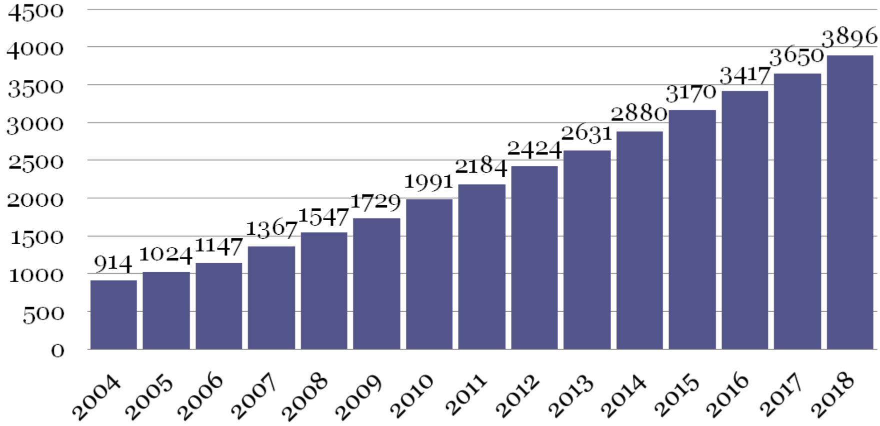

Figure 6 - Number of Internet Users (Source: Statista)

The rapid succession of digital disruptions calls for CIOs to create an organization that can rapidly adapt to new technologies and supports the new business models they enable. Cloud computing is still in the heart of the new environment. But Cloud computing is just one component of today's platform. Aligning business and IT objectives establish a collective vision as the basis for a digital business platform. By 2021, 40% of organizations will use enterprise architects to help ideate new business innovations made possible by emerging technologies. Digital platforms underpin some of the most successful business models, from Amazon to Uber. We see increasing interest from organizations in building a digital platform as a key part of their digital strategy. The platform is as much an architecture strategy as it is a business strategy. Leading Enterprise Architecture (EA) teams are fully involved in the platform's development, from ecosystems modeling and creating the business model, to the modular, service-oriented technical architectures that enable it.

Today a digital platform supports many digital businesses — the enterprise's own and those of its ecosystem partners. For instance, there is more computing power sitting idle on employees' desks and the device in their pockets than there is in all the world's data centers.

Can we speak about the perfect storm?

Possible yes. The "Holy Trinity" of digital disruption consists of blockchain, artificial intelligence (AI), and the Internet of Things. Blockchain's ability to support peer-to-peer microtransactions — combined with the intelligent decision-making ability of artificial intelligence (AI) and the sensory powers of IoT — will create never-before-possible businesses. AI applications including chatbots and virtual personal assistants are already being widely adopted across enterprises, and blockchain technologies are continuing to rapidly mature.

Data science can be split into the following units: a) applied machine learning and modeling, b) big data and data engineering, c) data monetization, d) data science in the Fourth industrial revolution, e) data analytics, f) deep learning and artificial intelligence, g) business intelligence, h) digital transformation, i) data science education, and j) open data.

In the following paragraphs, we can discuss the relationship between the Internet of Things and the mentioned disciplines.

MACHINE LEARNING AND THE INTERNET OF THINGS

Machine learning is a field of computer science and is a type of Artificial Intelligence (AI) that provides machines and a varied set of devices including the Internet of things with the ability to learn without explicit programming. Machine learning evolved from pattern recognition and computational learning theory. This chapter will briefly describe some essential concepts and algorithms necessary for the smart analysis of data. Table 2 presents some of the algorithms and data processing tasks that these algorithms perform. (Mohssen Mohammed, Muhammad Badruddin Khan, Eihab Bashier Mohammed Bashier, 2017)

Table 2. Overview of frequently used machine learning algorithms

K-Nearest Neighbors	Classification
Naive Bayes	Classification
Support Vector Machine	Classification
Linear Regression	Regression
Support Vector Regression	Regression
Classification and Regression Trees Classification	Regression
Random Forests Classification	Regression
Bagging Classification	Regression
K-means	Clustering
Density-Based Spatial Clustering of Applications with Noise	Clustering
Principal Component Analysis	Feature extraction
Canonical Correlation Analysis	Feature extraction
Feed Forward Neural Network	Regression/Classification/ Clustering/Feature extraction
One-class Support Vector Machines	Anomaly detection

Table 3 shows an overview of the machine learning algorithms for the Internet of Things cases.

Table 3. Overview of the machine learning algorithms to the Internet of Things cases

Machine learning Algorithm	IoT, Smart City use cases	Metric to Optimize References
Classification (Mohamed Amine Kafi, Yacine Challal, Djamel Djmpuri, Messaoud Doudou, Abdelmadjid Bouabdullah, Nadjib Badache, 2013)	Smart Traffic	Traffic Prediction, Increase Data Abbreviation
Clustering (YongruiQin, Quan Z.Sheng, Nickolas J.G.Falkner, Schahram Dustda, Hua Wang, Athanasios V.Vasilakos, 2016)	Smart traffic, Smart Health	Traffic Prediction, Increase Data Abbreviation
Anomaly Detection (Vikramaditya Yakkula, Diane Cook, 2010)	Smart Traffic, Smart Environment	Traffic Prediction, Increase Data Abbreviation, Finding Anomalies in Power Dataset
Support Vector Regression (P. Ni, C. Zhang, Y. Ji, 2014)	Smart Weather Prediction	Forecasting
Linear Regression (Wassim Derguech, Eanna Bruke, Edward Curry, 2014)	Economics, Market analysis, Energy usage	Real-Time Prediction, Reducing the Amount of Data
Classification and Regression Trees (Xiaolei Ma, Yao-Jan Wu, Yinhai Wang, Feng Chen, Jianfeng Liu, 2013)	Smart Citizens Real-Time Prediction	Passengers Travel Pattern
Support Vector Machine (Muhammad Aamir Khan, Aunsia Khan, Muhammad Nasir Khan, Sajid Anwar, 2014)	All Use Cases	Classify Data, Real-Time Prediction

K-Nearest Neighbors (Cao-Tri Do, Ahlame Douzal-Chouakria, Sylvain Marié, Michèle Rombaut, 2015)	Smart Citizen Passengers' Travel Pattern	The Efficiency of the Learned Metric
Naive Bayes Smart Agriculture (W. Han, Y. Gu, Y. Zhang, L. Zheng, 2014)	Smart Citizen Food Safety	Passengers Travel Pattern, Estimate the Numbers of Nodes
K-Means (Alberto M.C.Souza, Joseé R.A.Amazonas, 2015)	Smart City, Smart Home, Smart Citizen, Controlling Air and Traffic	Outlier Detection, fraud detection, Analyze Small Data set, Forecasting Energy Consumption, Passengers Travel Pattern, Stream Data Analyze
Density-Based Clustering (Muhammad Aamir Khan, Aunsia Khan, Muhammad Nasir Khan, Sajid Anwar, 2014)	Smart Citizen	Labeling Data, Fraud Detection, Passengers Travel Pattern
Feed Forward Neural Network (F. Ramalho, A. Neto, K. Santos, N. Agoulmine, et al., 2015)	Smart Health	Reducing Energy Consumption, Forecast the States of Elements, Overcome the Redundant Data and Information
Principal Component Analysis (Dorothy N. Monekosso, Paolo Remagnino, 2013)	Monitoring Public Places	Fault Detection
Canonical Correlation Analysis (Dorothy N. Monekosso, Paolo Remagnino, 2013)	Monitoring Public Places	Fault Detection
One-class Support Vector Machines (M. Shukla, Y. Kosta, P. Chauhan, 2015)	Smart Human Activity Control	Fraud Detection, Emerging Anomalies in the Data

From tables 2 and 3 we can see the following category in today machine learning scientists and engineers use:

- Classification
- Regression
- Combining models
- Clustering
- Feature extraction
- Neural networks
- Time series and sequence data
- Anomaly detection

Classification

K-Nearest Neighbors

The k-nearest neighbors' algorithm (k-NN), in pattern recognition, is a non-parametric method used for classification and regression. A commonly used distance metric for continuous variables is Euclidean distance. For discrete variables, (e.g. text classification) another metric can be used, such as the overlap metric (or Hamming distance). In k-NN regression, the k-NN algorithm is used for estimating continuous variables. Algorithms based on computing Euclidean or Mahanolobis distance. The form of the classification task is:

$$p(t = c|x, K) = \frac{1}{K} \sum_{iENk(x)} 1(ti = c)$$

$$y = \arg maxp(t = c|x.K) \tag{1}$$

The input vector x will be labeled by the mode of its neighbors' labels

KNN's main disadvantage of becoming significantly slower as the volume of data increases makes it an impractical choice in environments where predictions need to be made rapidly. Moreover, there are faster algorithms that can produce more accurate classification and regression results.

If you have sufficient computing resources to speedily handle the data you are using to make predictions, KNN can still be useful in solving problems that have solutions that depend on identifying similar objects. An example of this is using the KNN algorithm in recommender systems, an application of KNN-search. Implementation of KNN algorithm was shown in (Harrison, 2020).

Naive Bayes

Bayes theorem is one of the earliest probabilistic inference algorithms developed by Reverend Bayes (which he used to try and infer the existence of God no less) and still performs extremely well for certain use cases. It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as '*Naive*'.

Bayes formula is as follows:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \qquad (2)$$

Where A and B are events, and P(B)$\neq 0$ and
P(A|B) is a conditional probability: the likelihood of event A occurring given that B is true,
P(B|A) is also a conditional probability: the likelihood of event B occurring given that A is true
P(A) and P(B) are the probabilities of observing A and B respectively; they are known as the marginal probability

The term '*Naive*' in Naive Bayes comes from the fact that the algorithm considers the features that it is using to make the predictions to be independent of each other, which may not always be the case. So in our Diabetes example, we are considering only one feature, that is the test result. Say we added another feature, 'exercise'. Let's say this feature has a binary value of 0 and 1, where the former signifies that the individual exercises less than or equal to 2 days a week and the latter signifies that the individual exercises greater than or equal to 3 days a week. If we had to use both of these features, namely the test result and the value of the 'exercise' feature, to compute our final probabilities, Bayes' theorem would fail. Naive Bayes' is an extension of Bayes' theorem that assumes that all the features are independent of each other.

Support Vector Regression

Support Vector Machine can also be used as a regression method, maintaining all the main features that characterize the algorithm (maximal margin). The Support Vector Regression (SVR) uses the same principles as the SVM for classification, with only a few minor differences. First of all, because the output is a real number it becomes very difficult to predict the information at hand, which has infinite possibilities. In the case of regression, a margin of tolerance (epsilon) is set in approximation to the SVM which would have already been requested from the problem. But besides this fact, there is also a more complicated reason, the algorithm is more complicated therefore to be taken into consideration. However, the main idea is always the same: to minimize error, individualizing the hyperplane which maximizes the margin, keeping in mind that part of the error is tolerated.

COMBINING MODELS

Classification and Regression Trees (CART)

Classification and regression trees is a term used to describe decision tree algorithms that are used for classification and regression learning tasks.

The Classification and Regression Tree methodology, also known as the CART were introduced in 1984 by Leo Breiman, Jerome Friedman, Richard Olshen, and Charles Stone.

A classification tree is an algorithm where the target variable is fixed or categorical. The algorithm is then used to identify the “class” within which a target variable would most likely fall.

A regression tree refers to an algorithm where the target variable is and the algorithm is used to predict its value. As an example of a regression type problem, you may want to predict the selling prices of a residential house, which is a continuous dependent variable.

Classification trees are used when the dataset needs to be split into classes that belong to the response variable. In many cases, the classes are Yes or No. In other words, they are just two and mutually exclusive. In some cases, there may be more than two classes in which case a variant of the classification tree algorithm is used.

Regression trees, on the other hand, are used when the response variable is continuous. For instance, if the response variable is something like the price of a property or the temperature of the day, a regression tree is used. In other words, regression trees are used for prediction-type problems while classification trees are used for classification-type problems.

***

Random Forests and Bagging can not be described in this chapter.

CLUSTERING

K-Means

In the K-means algorithm, the objective is to cluster the unlabeled data set into K clusters (groups), where data points belonging to the same cluster must have some similarities. In the classical K-means algorithm, the distance between data points is the measure of similarity. Therefore, K-means seeks to find a set of K cluster centers, denoted as (s_1;...; s_k), such that the distances between data points and their nearest center are minimized

Time Series and Sequence Data

One of the “oldest” disciplines in these categories is time series. Sequential Data is any kind of data where the order matters. So we can assume that time series is a kind of sequential data, because the order matters. A time series is a sequence taken at successive equally spaced points in time and it is not the only case of sequential data. (Douglas Montgomery, Cheryl Jennings, Murat Kulahci, 2008), (Yafee, 1999) In the latter, the order is defined by the dimension of time. The first known examples are from financial mathematics. (Edgeworth, 1881). A sophisticated market analysis was presented in the book of Edgar Peters (Peters, 1994)

There are other cases of sequential data as data from text documents, (Petar Kocovic, Mutu Ramachandran at all, 2014) where you can take into account the order of the terms or biological data (DNA sequence, etc). The fact that you have sequential data is important for two reasons. First, you can into account the representation of the data, and also you can take it into account for the data modeling (e.g. Conditional Random Fields, Hidden Markov Models for text or genes, and ARIMA Models (Kocovic, 2015) for time series problems).

Neural Networks

One of the shortcomings of linear regression is that it is necessary to decide the types of basic functions. It is often difficult to choose the optimal basic functions. Therefore, in neural networks, we fix the number of basic functions but allow the model to learn the parameters of the basic functions. There exist many different types of neural networks, with different architectures, use cases, and applications. In subsequent subsections, we discuss the successful models used in smart data analysis. Note that neural networks are fast to process new data because they are compact models. However, in contrast, they usually require a large amount of computation to be trained. Moreover, they are easily adaptable to regression and classification problems.

BIG DATA AND THE INTERNET OF THINGS

When organizations are grabbing hold of the data for analysis purposes, IoT is acting as a major source for that data, and this is the point where the role of big data in IoT comes into the picture. Big data analytics is emerging as a key to analyzing IoT-generated data from "connected devices" which helps to take the initiative to improve decision making.

The role of big data in IoT is to process a large amount of data on a real-time basis and store them using different storage technologies.

IoT big data processing follows four sequential steps:

- A large amount of unstructured data is generated by IoT devices which are collected in the big data system. This IoT-generated big data largely depends on their 3V factors which are volume, velocity, and variety.

- In the big data system which is a shared distributed database, a huge amount of data is stored in big data files.
- Analyzing the stored IoT big data using analytic tools like Hadoop MapReduce or Spark
- Generating the reports of analyzed data.

It's not just that there is only interdependent relation between big data and IoT. As they help each other, in addition to that they hugely impact each other. The fact is the more the IoT grows it will place more demand on businesses regarding big data capabilities. For example, as IoT-generated data is increasing at a huge rate, conventional data storage technology is already being pushed to its limits. As a result, it demands more advanced and innovative storage solutions to handle these growing workloads resulting in updating the infrastructure of an organization's big data storage.

Similarly, the IoT big data combined applications accelerate the scope of research in both fields. So, IoT and big data technologies carry inter-dependency and need further development.

IOT PROJECTS THAT ANYONE CAN START

Small-scale projects are of interest to engineers. Here is the list of such projects that anyone can build for their purposes.

IoT based Weather Reporting System

To update the weather report manually is time-consuming. The necessity of an automated reporting update solution arises. IoT-based weather reporting system brings a solution where this system uses temperature, humidity rain sensors to monitor weather and report weather statistics online.

It works constantly and sends data via microcontroller to the webserver using the WIFI Internet. This system allows the user to set a threshold for a particular situation and alerts the user if weather reporting crosses the threshold value.

Important Features

- This system does not need human attention to monitor, as it is an automated system.
- It helps to collect data in tough environments like a volcano, minefield, polar zone, etc.
- Internet connection is needed for both ends.
- The future advancement of this project would be predicting weather forecasts and disasters.
- To build this project, you have to know how to use Arduino. (Arduino, 2020)
- It uses a learn-do-review methodology.

A further upgrade of this system is the Bi Data analysis collected from this weather system. This leads to a real meteo-system that can deliver weather forecasts, as well as historical data using smart weather prediction (see Table 3).

Touch-based Home Automation System

The Internet of things concept works with almost every machine. But when this is all about our home appliances, IoT proposes a smart, automated system. Using IoT-based automation systems users can control home stuff anywhere in the world.

In this article, we are talking about a touch-based home automation system based on a microcontroller. It contains WIFI, inbuilt touch sensing input pins, which makes it helpful to make IoT-based projects like this.

Important Features

- Adruino and ESP-32 microcontroller were used in this project.
- The user needs a smartphone or touchpad to control home kinds of stuff.
- Setup of ESP-32 in Adruino IDE.
- Minimum programming knowledge is needed.

A special case of this system is a smart garage opener. All in one built-in smartphone or tablet computer to open your garage replacing that clicker. Raspberry Pi is a standard component of the smart door opener.

A further improvement of such a system is based on the behavior estimating of the user(s) of such a system using K-means algorithms.

Facial Recognition Door Opener with Raspberry Pi

This is an intelligent Internet of things example that uses Raspberry Pi components (Raspberry Pi 4, 2020). Science fiction becoming true nowadays with the hand of IoT project ideas. A smart door secures the gateway and ensures the right person enters your home. Microsoft already made face API through their research.

Important Features

- This project is built on three phases that are 1. Data gathering 2. Training recognizer 3. Facial recognition.
- Python code is used for data gathering.
- Raspberry Pi camera is used for facial detection.
- This Internet of things example is not 100% accurate.

Support vector Machines and K-means represent tools for further improvements to this concept. Together with the programming language Python, this will be set for a fast and easy upgrade of the base concept.

Liquid Level Monitoring System

Liquid level monitoring system designed in a way that users can remotely check the level of the liquid. It has vast applications in the industrial sector where the user needs to monitor the level of liquid, whether it is below the mark of overflowing.

Theft detection, usage of chemicals, and leak detection are some of the usages of the liquid level monitoring system. Ultrasonic, Conductive, and float sensors are some of the few sensors for a monitoring system.

Important Features

- The ESP8266 Wifi module, produced by manufacturer Espresif Systems, (ESP8266EX, 2020) is used to connect with the Internet and transmit data to the required website.
- The ultrasonic float sensor will send data about the level of the liquid.
- Users will have a history of records from the website.

- It is easy to monitor from anywhere.

In the future, this Internet of things example will be developed for Bluetooth technology.

Further implementation of Linear regression and ARIMA algorithms can provide a forecast of the liquid level, which is necessary for some industrial applications.

IoT Based Air Pollution Monitoring System

Air pollution is a common problem nowadays. Different particles in the air like led, carbon dioxide, sulfur dioxide, pollen, and mold spores are making so much air pollution (PM2.5 and PM10 particles, HCHO and TVOC, as well as AQI index are part of the everyday life of the citizen). Air pollution can bring lots of diseases.

It is essential today to use a mechanism to measure air pollution in an area. Research on IoT projects brings a solution. Newly discovered IoT devices can monitor air pollution and save data to the web servers.

IoT project ideas like air pollution meters bring a solution to the existing problems like previous air pollution meter was out of memory after some time. But IoT device uses the Internet and saves data to the remote web server it has now become so easy to get a log of data within an area for specific days.

Important Features

This Internet of things example can also detect flammable gas leaks.

Particle matter detectors, gas sensors, temperature, and humidity sensors are used.

This type of project was built based on Arduino Uno.

Very helpful to detect air pollution close to the hospital or school.

This project is cost-efficient.

Various analytical tools are available for further processing the data, such are ARIMA, Linear regression, Clustering, and Support Vector Regression.

Night Patrolling Robot

Figure 7 – Night Patrolling Robot

Security is a common concern for all. As most of the crime occurs at night, the IoT project comes up with a solution that is a patrolling robot that uses a night vision camera. This robot patrols over a predefined path and detects alarming sounds.

If found, it scans the area with its 360 degrees moving camera and tries to detect any human face. Then it transmits the image to the nearby user who is running this whole IoT project. The user gets the alarming notification sent by the robot.

Important Features

- IR sensor makes it happen that a robot will patrol through a defined path.
- USB camera and Raspberry Pi are connected.
- Python's language helps make its software.
- This robot can be cheaper than hiring multiple security guards.

Smart Parking System

Figure 8 - Smart parking concept

Figure 8.1 – Sensor for a free parking place

Figure 8.2 – Wi-Fi antenna for collecting the data from the sensors

Finding a parking space is a problem for the driver. Sometimes it wastes a lot of time for the driver to find a parking space. IoT-based project smart parking system brings a solution. A major purpose of this project is to avoid unnecessary traveling by a driver for the parking area. Monitoring the whole area at the run time gives the driver an image of the entire parking area, and the user can select that free parking space.

Important Features

- Daily life problem solution by IoT-based ideas.
- It uses an IR sensor for detecting free parking space.
- Illegal parking can be reduced with this Internet of things example.
- This IoT-based project is built on the Arduino board.
- The infrared proximity sensor can be used for finding space.

The extension of this concept leads to the further development of smart vehicles. In these concepts, there is an option for communication between cars and the parking system. Various parking systems in one area can offer different prices for parking, and cars can decide to change the parking lot (with/without communication with the owner of the car). Further software for Ai can be implemented in this concept.

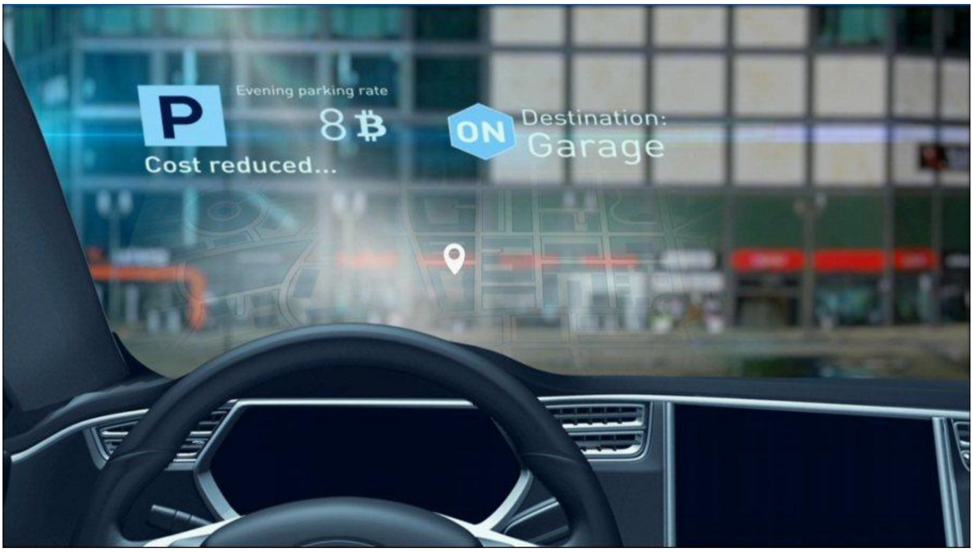

Figure 9 – Decision-making system in an autonomous vehicle for changing parking places.

IoT Based Health Monitoring System

Health is the most valuable asset of human life. Over time the life of people become so stressful that we are giving less care to our health. People are hardly going for a checkup. IoT projects like health monitoring systems can make a solution here with the devices that monitor our health regularly and send data to the doctor.

The doctor can check the current situation at any time and anywhere in the world. Sensors in the body of the patient can estimate blood pressure, sugar level, and heartbeat and immediately alarm the doctor if it is higher than normal.

Important Features

- Adruino will generate the output.
- The doctor can check the current condition of the patient using his smartphone.
- Use of firebase for run-time data.

Communication between Arduino and the Android app is needed.
This Internet of things example uses Bluetooth technology.

5G EDGE or FOG is a constitutive part of this system

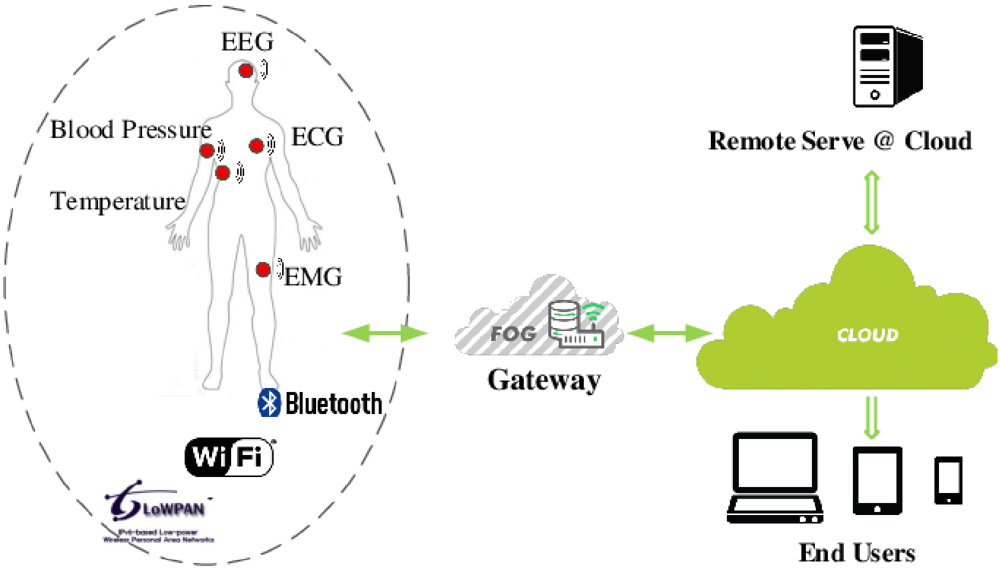

Figure 10 – IoT Based Health Monitoring System

Clustering and Feed-Forward neural networks are further improvements of this concept.

IoT Based Traffic Management System

Almost every metropolitan city faces traffic problems. As the population is increasing day by day, the necessity of a smart traffic management system is undeniable. IoT-based projects like smart traffic management systems can reduce the traffic problem. Sometimes it is difficult for the ambulance to cross the traffic.

Counting the advantage of an ambulance, this management system connects with the ambulance driver and helps to find the signal where traffic flow can be controlled dynamically. This Internet of things example also monitors traffic rules violators.

Important Features

- Finding an emergency path for an emergency is easy.
- This IoT project can be used anywhere.
- Can identify traffic violators at night.

- It shows green light only for an ambulance, fire truck, or emergency vehicle.

Classification, clustering, and anomaly detection are some algorithms that can improve traffic monitoring and control.

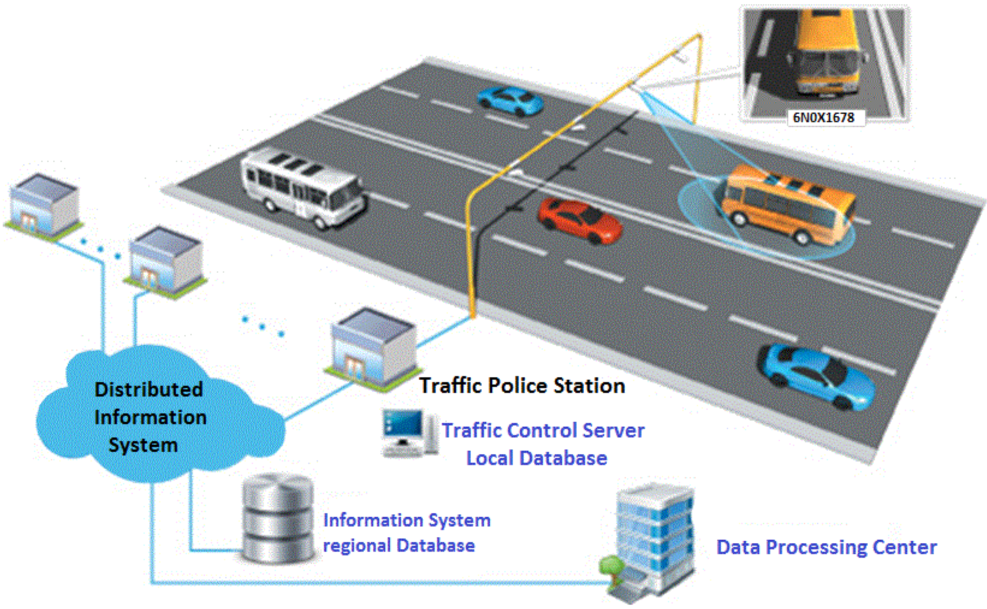

Figure 11 – Traffic Management System

IoT Based Garbage Monitoring System

This IoT project idea uses an ultrasonic sensor to detect the level of garbage in each garbage bin and send those data to the main IoT program. A webpage shows the level of garbage on each garbage bin and highlights the amount of garbage on each bin. A buzzer puts on when garbage is over the limit. One of the first solutions that were implemented was in the Municipality of Dubai where tens of companies compete for garbage collecting. (Petar Kocovic, Vojkan Vaskovic, 2014)

Important Features

- For sending data, each bin uses Wifi Modem.

IoT gecko web development platform used for graphical representation.
Programming language C is used here.

Adruino Compiler and HC-SR04 ultrasonic sensors can be used for this.

FIGURE 12 – RFID based recycle bin and container management system

IoT Based Street Light Monitoring System

Artificial light sources play an indispensable role in the daily life of any human being. Electrical light sources are responsible for the energy consumption of around 1/6 to 1/5 of worldwide electricity production. (Zissis, 2016) IoT project ideas like street light monitoring systems bringing a solution here. The project consists of smart street lights that have sensors to monitor humans or vehicles around.

Upon sensing the movement, the sensor sends data to the microcontroller; then it turns the street light on. If there is no movement microcontroller makes the switch off. That's how this IoT project saves energy. Checking faulty street lights is another advantage of this Internet of things example.

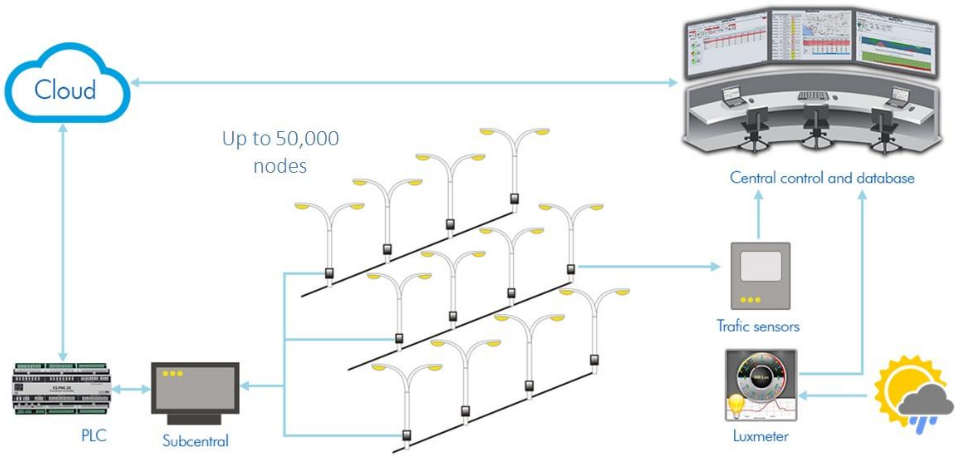

Figure 13 – Smart lighting

CONCLUSION

Data Analytics has a significant role to play in the growth and success of IoT applications and investments. Analytics tools will allow the business units to make effective use of their datasets as explained in the points listed below.

Volume: There are huge clusters of data sets that IoT applications make use of. Business organizations need to manage these large volumes of data and need to analyze the same for extracting relevant patterns. These datasets along with real-time data can be analyzed easily and efficiently with data analytics software.

Structure: IoT applications involve data sets that may have a varied structure as unstructured, semi-structured, and structured data sets. There may also be a significant difference in data formats and types. Data analytics will allow the business executive to analyze all of these varying sets of data using automated tools and software.

Driving Revenue: The use of data analytics in IoT investments will allow the business units to gain insight into customer preferences and choices. This would lead to the development of services and offers as per the customer demands and expectations. This, in turn, will improve the revenues and profits earned by the organizations.

Competitive Edge: IoT is a buzzword in the current era of technology and there are numerous IoT application developers and providers present in the market. The use of data analytics in IoT investments will provide a business unit to offer better services and will, therefore, provide the ability to gain a competitive edge in the market.

REFERENCES

The Constrained Application Protocol (CoAP). (2020, January 2020). Retrieved from IETF: https://datatracker.ietf.org/doc/rfc7252/

1888.3-2013 - IEEE Standard for Ubiquitous Green Community Control Network: Security. (2020, January 18). Retrieved from IEEE Standards Association: https://standards.ieee.org/standard/1888_3-2013.html

(2020, January 2). Retrieved from MQTT: mqtt.org

(2020, January 2). Retrieved from Github: https://github.com/darconeous/old-coap-archived/tree/smcp-0.6

(2020). *January* (p. DDIS Foundation). 2: https://www.dds-foundation.org/.

(2020, January 2). Retrieved from Websocket: https://websocket.org/

(2020, January 22). Retrieved from Oma SpecWorks: https://www.omaspecworks.org/

(2020, January 27). Retrieved from Arduino: https://www.arduino.cc/

2413-2019 - IEEE Approved Draft Standard for an Architectural Framework for the Internet of Things (IoT). (2020, January 18). Retrieved from IEEE SA: https://standards.ieee.org/content/ieee-standards/en/standard/2413-2019.html

About nanoIP. (2020, January 2). Retrieved from nanoIP: http://www.cwc.oulu.fi/nanoip/

About UPnP. (2020, January 2). Retrieved from Open Connectivity Foundation: https://openconnectivity.org/developer/specifications/upnp-resources/upnp/

Alberto M.C.Souza, Joseé R.A.Amazonas. (2015). An Outlier Detect Algorithm using Big Data Processing and Internet of Things Architecture. *Procedia Computer Science* (pp. 1010-1015). Elsevier.

AMQP. (2020, January 2). Retrieved from http://www.amqp.org/

ANT. (2020, January 2). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/ANT_(network)

Bluetooth. (2020, January 2020). Retrieved from Bluetooth: https://www.bluetooth.com/

C. Jennings, Z. Shelby, J. Arkko. (2020, January 18). *Media Types for Sensor Markup Language (SENML)*. Retrieved from Network Working Group: https://tools.ietf.org/html/draft-jennings-senml-08

Cao-Tri Do, Ahlame Douzal-Chouakria, Sylvain Marié, Michèle Rombaut. (2015). Multiple metric learning for large margin kNN classification of time series. *Signal Processing Conference (EUSIPCO)* (pp. 2346-2350). IEEE.

Carsten Bormann, Geoffrey Mulligan, Ted Lemon. (2020, January 19). *IPv6 over Low power WPAN (6lowpan)*. Retrieved from IETF: https://datatracker.ietf.org/wg/6lowpan/charter/

Carsten Bormann, Jaime Jimenez, Alexey Melnikov. (2020, January 19). *Constrained RESTful Environments (core)*. Retrieved from IETF: https://datatracker.ietf.org/wg/core/charter/

Control Your World. (2020, January 2). Retrieved from Zigbee Alliance: https://zigbeealliance.org/

DDS Standard. (2020, January 19). Retrieved from DDS: https://www.dds-foundation.org/

Dorothy N. Monekosso, Paolo Remagnino. (2013, June 15). Data reconciliation in a smart home sensor network. *Experts Systems with Applications*, pp. 3248-3255.

Douglas Montgomery, Cheryl Jennings, Murat Kulahci. (2008). *Introduction to Time Series Analysis and Forecasting.* Jon Wiley & Sons.

E. Rescorla, N. Modadugu. (2020, January 2). *Datagram transport layer Security*. Retrieved from Stanford University: https://tools.ietf.org/html/rfc4347

Eclipse Mosquitto. (2020, January 2). Retrieved from Mosquitto Eclipse Foundation: http://mosquitto.org/

Edgeworth, F. (1881). *Mathematical Physics.* London: C. KEGAN PAUL & CO., 1 PATERNOSTER SQUARE.

Edystone. (2020, January 2). Retrieved from GitHub: https://github.com/google/eddystone

EPCglobal. (2020, January 20). Retrieved from GS1: https://www.gs1.org/epcglobal

Ericsson and Orange in the Internet of Things trial with EC-GSM-IoT. (2020, February 3). Retrieved from Ericsson: https://www.ericsson.com/en/press-releases/2016/2/ericsson-and-orange-in-internet-of-things-trial-with-ec-gsm-IoT

ESP8266EX. (2020, January 27). Retrieved from Espressif: https://www.espressif.com/en/products/hardware/esp8266ex/overview

ETSI. (2020, January 19). Retrieved from ETSI: https://www.etsi.org/

F. Ramalho, A. Neto, K. Santos, N. Agoulmine, et al. (2015). Enhancing eHealth smart applications: A Fog-enabled approach. *17th International Conference on e-health Networking, Applications and Services (HealthCom)* (pp. 323-328). IEEE.

Harrison, O. (2020, January 26). *Machine Learning Basics with the K-Nearest Neighbors Algorithm*. Retrieved from Towards Data Science:

https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761

Host Identity Protocol (HIP). (2020, January 20). Retrieved from IETF: https://www.ietf.org/proceedings/60/264.htm

I.I. (2005). *Internet of Things.* Geneva: ITU.

IEEE 1377:2012 - Standard for Utility Industry Metering Communication Protocol Application Layer (End Device Data Tables). (2020, January 20). Retrieved from IEEE Standards Association: https://standards.ieee.org/standard/1377-2012.html

IEEE 1905.1-2013 - IEEE Standard for a Convergent Digital Home Network for Heterogeneous Technologies. (2020, January 18). Retrieved from IEEE Standards Association: https://standards.ieee.org/standard/1905_1-2013.html

IEEE 802.15 WPAN™ Task Group 4 (TG4). (2020, January 2). Retrieved from IEEE 802: http://www.ieee802.org/15/pub/TG4.html

IEEE 802.16p-2012 - Standard for Air Interface for Broadband Wireless Access Systems--Amendment 1: Enhancements to Support Machine-to-Machine Applications. (2020, January 18). Retrieved from IEEE Standards Association: https://standards.ieee.org/standard/802_16p-2012.html

IIC. (2020, January 22). Retrieved from https://www.iiconsortium.org/

Internet of Things. (2020, January 2). Retrieved from British Standards Institution: https://www.bsigroup.com/en-GB/industries-and-sectors/Internet-of-Things/

Internet of Things. (2020, January 2020). Retrieved from IEEE Standards Institute: https://standards.ieee.org/initiatives/IoT/index.html

Internet of Things Global Standards Initiative. (2020, January 19). Retrieved from ITU: https://www.itu.int/en/ITU-T/gsi/IoT/Pages/default.aspx

IOTDB. (2020, January 18). Retrieved from http://IoTdb.org/

IoTivity. (2020, January 18). Retrieved from IoTivity: https://IoTivity.org/

IPv6. (2020, January 2020). Retrieved from IETF: www.ietf.org

IPv6 over Low power WPAN (6lowpan). (2020, January 1). Retrieved from IETF: https://datatracker.ietf.org/wg/6lowpan/documents/

ISA. (2020, January 19). Retrieved from https://www.isa.org/

ISO. (2020, January 20). Retrieved from ISO: https://www.iso.org/technical-committees.html

ITU. (2020, January 1). *Internet of Things Global Standards Initiative*. Retrieved from ITU: https://www.itu.int/en/ITU-T/gsi/IoT/Pages/default.aspx

J. Iyengar, M. Thomson. (2020, January 1). *QUIC: A UDP-Based Multiplexed and Secure Transport*. Retrieved from IETF: tools.ietf.org/html/draft-ietf-quic-transport-22#section-1

John Mark, Paul Hufnagel. (2020, January 18). *The IEEE 1451.4 Standard for Smart Transducers*. Retrieved from IEEE Standards Association: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/tutorials/1451d4.pdf

Kocovic, V. (2015, July). ARIMA model and its usage in Industry. *Master work*. Belgrade: University of Belgrade. School of Mechanical Engineering.

Latest SOAP versions. (2020, January 2). Retrieved from W3C: https://www.w3.org/TR/soap/

LEMONBEAT OS. (2020, January 18). Retrieved from Lemonsbeat: https://www.lemonbeat.com/lemonbeat-IoT-operating-system/

LLAP protocol spec #6. (2020). *January* (p. Github). 2: https://github.com/CisecoPlc/LLAPSerial/issues/6.

Long Term Evolution for Machines: LTE-M. (2020, January 3). Retrieved from GSMA: https://www.gsma.com/IoT/long-term-evolution-machine-type-communication-lte-mtc-cat-m1/

LWM2M – M2M The Lighter Way. (2020). *January* (p. Y.I Readings). 2: https://yucianga.info/?p=786.

M. Belshe, R. Peon, M. Thomson. (2020). HTTP. *January* (p. Hypertext Transfer Protocol Version 2 (HTTP/2)). 2: https://httpwg.org/specs/rfc7540.html.

M. Pei, A. Atyeo, N. Cook, M. Yoo, H. Tschofenig. (2020, January 18). t*he Open Trust Protocol (OTrP)*. Retrieved from Internet engineering task Force: https://tools.ietf.org/id/draft-ietf-teep-opentrustprotocol-00.html

M. Shukla, Y. Kosta, P. Chauhan. (2015). Analysis and evaluation of outlier detection algorithms in data streams. *International Conference on Computer, Communication, and Control (IC4)* (pp. 1-8). IEEE.

Mihini/M3DA Specification. (2020, January 2). Retrieved from ECLIPSE Foundation: https://wiki.eclipse.org/Mihini/M3DA_Specification

Mohamed Amine Kafi, Yacine Challal, Djamel Djmpuri, Messaoud Doudou, Abdelmadjid Bouabdullah, Nadjib Badache. (2013). A Study of Wireless Sensor Networks for Urban Traffic Monitoring: Applications and Architectures. *The 4thInternational Conference on Ambient Systems, Networks and Technologies* (pp. 617-626). Elsevier.

Mohssen Mohammed, Muhammad Badruddin Khan, Eihab Bashier Mohammed Bashier. (2017). *Machine Learning - Algorithms and Application*. Boca Raton, FL: CRC Press.

Muhammad Aamir Khan, Aunsia Khan, Muhammad Nasir Khan, Sajid Anwar. (2014). A novel learning method to classify data streams in the internet of things. *Software Engineering Conference (NSEC)* (pp. 61-66). IEEE.

Muscanello, L. (2020, January 2). *CiCN*. Retrieved from https://wiki.fd.io/view/Cicn

Narrowband – Internet of Things (NB-IoT). (2020, January 3). Retrieved from the Internet of Things: https://www.gsma.com/IoT/narrow-band-Internet-of-things-nb-IoT/

NFC Forum. (2020, January 2). Retrieved from https://nfc-forum.org/: https://nfc-forum.org/

OASIS Message Queuing Telemetry Transport (MQTT) TC. (2020, January 19). Retrieved from OASIS: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=mqtt

Object Name Service (ONS). (2020). *January* (p. GS1). 2: https://www.gs1.org/standards/epcis/epcis-ons/2-0-1.

OCF. (2020, January 22). Retrieved from https://openconnectivity.org/allseen-alliance-merges-open-connectivity-foundation-accelerate-Internet-things/

OCF SOLVING THE IOT STANDARDS GAP. (2020, January 18). Retrieved from OCF: https://openconnectivity.org/

OMG. (2020, January 19). Retrieved from OMG: https://www.omg.org/

OpenWSN. (2020, January 22). Retrieved from https://openwsn.atlassian.net/wiki/spaces/OW/overview

OSIOT. (2020, January 9). Retrieved from http://osIoT.org/

P. Ni, C. Zhang, Y. Ji. (2014). A hybrid method for short-term sensor data forecasting in the Internet of Things. *11th International Conference on Fuzzy Systems and*. IEEE.

Paho. (2020, January 20). Retrieved from Eclipse: https://www.eclipse.org/proposals/technology.paho/

Petar Kocovic, Mutu Ramachandran at all. (2014, April 17). EMOTIONALLY ANNOTATED DATABA. *Horizon 2020 - H2020-ICT-2014-1*. European Commission.

Petar Kocovic, Vojkan Vaskovic. (2014). IT in New Information Economy. *Proceedings "Managing and Development of Regions of the Eurasian Space: Status, Problems, and Prospects*, (pp. 131-134). Belgrade.

Peters, E. (1994). *Fractal Market Analysis*. John Wiley.

Postel, J. (2020, January 1). *User Datagram protocol*. Retrieved from RFC 768: https://tools.ietf.org/html/rfc768

Products Tagged "DigiMesh". (2020, January 2). Retrieved from Digi: https://www.digi.com/products/tag/digimesh

RAML. (2020, January 18). Retrieved from RAML: https://raml.org/

Raspberry Pi 4. (2020, January 27). Retrieved from https://www.raspberrypi.org/

Reactive Streams. (2020). *January* (p. Reactive Streams). 2: http://www.reactive-streams.org/.

Representational state transfer. (2020). *January* (p. Wikipedia). 2: https://en.wikipedia.org/wiki/Representational_state_transfer.

Routing Over Low power and Lossy networks (roll). (2020, January 2). Retrieved from IETF: https://datatracker.ietf.org/wg/roll/documents/

Routing Over Low power and Lossy networks (roll). (2020, January 19). Retrieved from IETF: https://datatracker.ietf.org/wg/roll/charter/

S. Cheshire, M. Krochmal. (2020, January 2). *Multicast DNS*. Retrieved from IETF: https://tools.ietf.org/html/rfc6762

Sade Kuyoro, Folasade Osisanwo, Omoyele Akinsowon. (2015). Internet of Things (IoT): An Overview. *3rd International Conference on Advances in Engineering Sciences & Applied Mathematics (ICAESAM'2015) March 23-24*, (pp. 53-58). London.

Self-powered Wireless Standard. (2020, January 2). Retrieved from EnOcea: https://www.enocean.com/en/

Semantic Sensor Net Ontology. (2020, January 18). Retrieved from W3C: https://www.w3.org/2005/Incubator/ssn/wiki/Semantic_Sensor_Net_Ontology

Semantic Sensor Net Ontology. (2020, January 19). Retrieved from https://www.w3.org/2005/Incubator/ssn/wiki/Semantic_Sensor_Net_Ontology

Sensor Model Language (SensorML). (2020, January 18). Retrieved from OGC: https://www.opengeospatial.org/standards/sensorml

Simple Sensor Interface protocol. (2020). *January* (p. Wikipedia). 2: https://en.wikipedia.org/wiki/Simple_Sensor_Interface_protocol.

Standards for M2M and the Internet of Things. (2020, January 19). Retrieved from ONEM2M: http://www.onem2m.org/

Statista. (2020, January 22). *Half of the World will be on the World Wide Web*. Retrieved from www.statista.com

Statista. (2020, January 1). *Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025*. Retrieved from Statista: https://www.statista.com/statistics/471264/IoT-number-of-connected-devices-worldwide/

STOMP Servers. (2020, January 2). Retrieved from STOMP: https://stomp.github.io/implementations.html

Tara Salman, Raj Jain. (2017, March). A Survey of Protocols and Standards for Internet of Things. *Advanced Computing and Communications*, pp. Vol 1, No 1.

The universal messaging standard. (2020, January 2). Retrieved from XMPP: https://xmpp.org/

uIP (Micro IP). (2020, January 1). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/UIP_%28micro_IP%29

Vavra, B. (2020, January 2). *ISA100.11a standard closes in on approval*. Retrieved from Plant Engineering: https://www.plantengineering.com/articles/isa100-11a-standard-closes-in-on-approval/

Vidales, M. (2020, January 2). *802.15.4 Wireless for Internet of Things Developers*. Retrieved from https://blog.helium.com/802-15-4-wireless-for-Internet-of-things-developers-1948fc313b2e

Vikramaditya Yakkula, Diane Cook. (2010). Outlier Detection in Smart Environment Structured Power Datasets. *Sixth International Conference on Intelligent Environments (IE),* (pp. 29-33). IEEE.

W. Han, Y. Gu, Y. Zhang, L. Zheng. (2014). The data-driven quantitative trust model for the Internet of agricultural things. *International Conference on the Internet of Things* (pp. 31-36). IEEE.

Wassim Derguech, Eanna Bruke, Edward Curry. (2014). An Autonomic Approach to Real-Time Predictive Analytics Using Open Data and Internet of Things Autonomic and Trusted Computing, and IEEE 14th Intl Conf on Scalable Computing and Communications and its Associated Workshops (UTC-ATCScalCom). *11th Intl Conf on and IEEE 11th Intl Conf on and* (pp. 204-211). IEEE.

WEB OF THINGS COMMUNITY GROUP. (2020, January 19). Retrieved from https://www.postscapes.com/Internet-of-things-protocols/

What is the LoRaWAN® Specification? (2020, January 3). Retrieved from LoRA Alliance: https://lora-alliance.org/about-lorawan

WiMax Forum. (2020, January 2). Retrieved from WiMax Forum: http://wimaxforum.org/home

X.509. (2020, January 18). *X.509: Information technology - Open Systems Interconnection - The Directory: Public-key and attribute certificate frameworks*. Retrieved from ITU: https://www.itu.int/rec/T-REC-X.509

Xiaolei Ma, Yao-Jan Wu, Yinhai Wang, Feng Chen, Jianfeng Liu. (2013, November 13). Mining smart card data for transit riders' travel patterns. *Transportation Research Part C: Emerging Technologies* (pp. 1-12). Elsevier.

XMPP-IoT. (2020, January 2). Retrieved from Github: https://github.com/joachimlindborg/XMPP-IoT

Yafee, R. (1999). *Introduction to Time Series and Forecasting*. Academic Press.

YongruiQin, Quan Z.Sheng, Nickolas J.G.Falkner, Schahram Dustda, Hua Wang, Athanasios V.Vasilakos. (2016). When things matter: A survey on data-centric Internet of things. *Journal of Network and Computer Applications*, 137-153.

Zissis, G. (2016). Energy Consumption and Environmental and Economic Impact of Lighting: The Current Situation. In *Handbook of Advanced Lighting Technology* (pp. 1-13). Springer Link.