Mindtree
Welcome to possible

TECH BEACON

THE TECHNOLOGY HARBINGER

TECH BEACON

THE TECHNOLOGY HARBINGER

Technology is core to Mindtree. It is extremely important in building capability and capacity for us to keep a sharp focus on emerging technologies. We also make it a point to understand the application of these technologies in the context of our customers. In this effort, we constantly explore and experiment with emerging technologies and gauge their maturity levels for consumption by enterprises. We do this by comparing and contrasting them with similar tools already in use.

Tech Beacon is an initiative to compile outcomes of our experimentation and technology adoption guidance for enterprises. We group technologies under three categories: Invest, Experiment and Watch. We explain these categories in detail in this report.

Direction

Tech Beacon provides direction to enterprises on various technologies by categorizing them as Invest, Experiment or Watch technologies.

Invest

These technologies have a high potential for generating business value in the near future. Mindtree intends to build capability internally or identify a partner, who can help our project teams learn to bootstrap these technologies. We arrive at these technologies by observing customer demands, industry trends, opportunities and open-source community support. Most of these technologies were either implemented in production-grade applications for customers or in internal reference implementations.

Experiment

These technologies have not yet seen mainstream adoption but show huge promise for the near future. We arrive at these technologies by listening to our customers' interests, following industry buzz, and noting indications of increased activity in various forums and analyst reports. We need to experiment and build capability in these technologies to have an early adopter edge.

Watch

The technologies in this category show promise through their architecture, the business cases they support, and a community push for adoption. These technologies are not yet adopted by enterprises, but are being evaluated by enterprises for their adoption benefits. While we don't have to invest much in them yet, we need to keep a close watch on these technologies and move them to the "Experiment" category as they mature or become more relevant to enterprises.

Table of
Contents

Architecture

Business needs are rapidly evolving and are driving architecture evolution. In this section, we present certain key business needs, and the architecture paradigms that have come about to support these business needs.

Need for Speed: Micro-services, DevOps

Cut-throat competition and a need to keep pace with ever-changing customer desires make businesses ask for high agility within their IT systems with the ever-changing business. The architecture style that supports this agility is of Micro-services. Micro-services, often called "SOA done right", is about building smaller, decoupled subsystems, that can evolve rapidly and independently. Micro-services, supported with DevOps – the mindset and processes to quickly effect changes, enables business systems to be nimble.

Always Open for Business: Resiliency Engineering

Globalization requires IT systems to be always on, and be always ready for customers and employees across the world. Daily scheduled downtimes are no longer viable, and even small outages make headlines. Resiliency Engineering, to ensure that systems are never down, even under heavy load, requires that systems are architected for availability, fault tolerance, and self-healing, at every layer. Beyond redundancy, multiple mechanisms such as resource management, circuit breakers, and deep instrumentation and monitors are all part of building continuously available systems.

Customer Intimacy: Big Data Analytics, NLP, ML

The differentiator for business today is no longer only services, but services that are highly personalized. The ability to harvest data from all possible sources and building a rich customer profile is critical to supporting such personalization. Building such a 360° view of the customer requires a single version of the customer, synthesized from many different sources: Both structured (such as CRM systems, and systems of record), and unstructured (such as social media interactions, and descriptive feedback). This requires the architecture constructs of big data storage and analytics, text mining and NLP and Machine Learning for segmentation. Such a view of the customer is surfaced by API that allow consumption for downstream personalization systems.
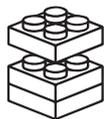
All Partners are Welcome: API Monetization

Partners have always been an essential part of business, but they have now evolved from a supporting role, to become primary sales channels. To support diverse partners to sell, enforcing a single storefront is not viable, and so the unit of interaction is the API that enables partners to dip into catalogs and to create orders. API, viable; so, to be written with a view to be able to have them monetized and metered.

The Human Touch: Conversational Apps, AR and VR

The bid to make UI simpler and ever more intuitive eventually ends at being able to interact using our core interactions of speech (or text) and vision. Coupled with app fatigue, conversational apps are the new wave of interaction. Architecturally, this translates to a Conversational Layer as a channel, in addition to the traditional Web and mobile channels, amply supported by the requisite API. This layer includes NLP and speech to text conversions, to be able to understand people as they speak/type. Similarly, using Augmented Reality and Virtual reality to meld physical entities and information systems, is a new way of interaction.
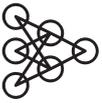
A Thousand Words: Visual Apps, Deep Learning

Having a picture and saying "It looks like that", is much easier for people, than poring through deeply nested menus of categories and subcategories of products. This leads to Visual Apps, which allow image-based searches. Image recognition using Deep Learning, to ensure recognition in the wild, is a new way of interaction and search.

Split-second Reaction Time: Lambda and Kappa

Businesses want to be able to react to a change in customer context "then and there." It is no longer viable to update a profile after 24 hours to be ready for the next interaction; it is the current interaction that matters. The architecture to support this is the Lambda Architecture, which consists of a speed layer for stream-based interactions and a batch layer to update based on the universe of data. Kappa Architecture, which leverages Spark or equivalent, moves this to the "process and store" paradigm.

Everything Web Scale: Reactive, Actors

With numerous channels, many partners and global customer presence, the challenge becomes that of scaling systems, so, they are always ready for more business. Scale-out paradigms, coupled with Reactive Programming at all layers, ensure that individual components don't become a bottleneck. Actor Systems are a way of building such reactive systems.

No More Servers: Server-less Architecture

With the advent of cloud, businesses have reduced their dependency on prolonged hardware procurement cycles and are able to provision servers on the go. Taking this further, not having even that step of standing up a server, loading up an image, and wiring it up, make systems even more quick to set up. This has led to Server-less Architecture, wherein the landscape is built on services that are managed, not VMs – for example, AWS Lambda or Azure functions, managed cache services, and managed Spark or Hadoop instances, among many others.

The Magic Formula: Algorithm Economy

While data is key and is closely guarded, having a lot of data, by itself, is no good, unless business has a way of understanding this data. Algorithms to understand this data, segment customers, infer preferences, and predict outcomes are as essential. The Algorithm Economy, where algorithms become the differentiator are the new wave. Beyond statistical models, machine learning and deep learning models are the IP that is closely guarded as the secret sauce.

Focus Areas

Mindtree presents a point of view on each of the following broad areas of focus in the technology domain.
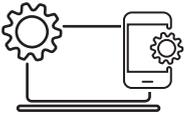
Application Development

The open-source market has never been more vibrant. Today, open-source technologies drive entire technology domains such as Web application and application programming interface (API) development frameworks, big data and analytics technologies. Commercial software vendors are also making available a plethora of new technology products.

Internet of Things

The Internet of Things continues to emerge as a major focus across most industry sectors. With the huge market potential of IoT, the need for standardization in the near term is real. Major IT solution providers have jumped in and started providing a variety of tools and technologies.

Cognitive Computing

Cognitive computing enables natural and contextual interaction with tools that augment users' experiences, using the power of machine-based intelligence. Tap into an ever-growing collection of powerful artificial intelligence algorithms for vision, speech, language and knowledge.

Blockchain

Blockchain is emerging as a fundamental technology that can potentially have a disruptive impact on traditional business models. It provides an incorruptible digital ledger of economic transactions that can be programmed to record not just financial transactions but virtually everything of value.
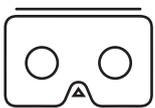
VR/AR

Virtual reality started luring many organizations this year across the world, with the launch of many cheaper variants of the headwear; expanding its niche market.
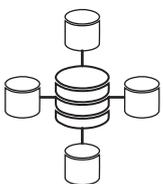
Big Data & Analytics

Data is a reflection of the organizations' maturity to measure processes, models and outcomes. In today's world, data imprints are growing and, hence, Big Data and Analytics have become the backbone of business processes driving personalization, automation and business insights. This is often used in conjunction with fast data, machine learning and parallel processing at scale.

Application Development

Invest

- C#7.0
- .Net Core 1.0
- Angular 2.0
- Apache Mesos
- ASP.NET Core 1.0
- Document DB
- Kubernetes
- MEAN
- MS SQL Server 2016
- ReactJS
- Redis
- Scala
- Spring Boot
- Spring Cloud
- Typescript
- WebRTC
- xUnit.net

Experiment

- C#7.0
- Consul by Hashicorp
- Go
- Java 9
- Microsoft Nano Server
- Microsoft Teams
- Microsoft Visual Studio 2017
- MS ASP.NET WebHooks
- MS SharePoint and Office Apps
- Project Orleans
- Vault by Hashicorp

Watch

- .Net Core 2.0
- Kitura
- Power BI
- Swift on Server
- Universal Windows Platform
- Windows Container

$ Invest

.Net Core 1.0

.NET Core is a general purpose, modular, cross-platform and open-source implementation of .NET. It includes a runtime, framework libraries, compilers and tools that support a variety of processor and OS targets. These components can be used together or separately. .NET Core 1.0 is optimized for server workloads and provides a subset of the APIs that are available within the .NET Framework.

Angular 2.0

AngularJS is a structural framework for creating dynamic web apps. It lets you use HTML as the template language and lets you extend HTML's syntax to express your application's components clearly and succinctly. AngularJS's data binding and dependency injection eliminate much of the code you would otherwise have to write.

Apache Mesos

Apache Mesos is a cluster manager that provides efficient resource isolation and sharing across distributed applications or frameworks. It abstracts CPU, memory, storage, and other compute resources away from machines (physical or virtual), enabling fault-tolerant and elastic distributed systems to easily be built and run effectively. Mesos is open-source software originally developed at the University of California at Berkeley.

ASP.NET Core 1.0

ASP.NET Core 1.0 is a significant redesign of ASP.Net, and is a new open-source and cross-platform framework for building modern cloud-based Web applications using .NET. It is a complete rewrite that unites the previously separate ASP.NET MVC and Web API into a single programming model.

Document DB

Document DB is a Microsoft Azure offering for NoSQL database as a service. It is a highly scalable schema-free database, having rich query capabilities and transaction processing using a subset of SQL. It also provides a server side programming model, where stored procedures, triggers and user-defined functions can be written using JavaScript. This provides a way to update multiple documents with atomic transactions. Unlike other NoSQL databases, it indexes every property in a document as soon as it is added and also gives a way of tuning your indexing based on storage and performance requirements. DocumentDB also allows the user

to tune its consistency using any of four choices: strong, eventual, bounded staleness and session. There is also an option to scale the performance up or down by choosing one of the performance tiers for a particular collection.

Kubernetes

Kubernetes is an open-source system for automating deployment, scaling, and management of containerized applications. It groups containers that make up an application into logical units for easy management and discovery.

MEAN

MEAN is a framework for building MongoDB, Node.js, Express and AngularJS based applications. It is designed to provide a quick and organized way to start developing MEAN-based web apps with useful modules like Mongoose and Passport. These and other such modules come pre-bundled and configured along with this framework. It mainly tries to address the connection points between existing popular frameworks and solve common integration problems.

MS SQL Server 2016

MS SQL Server 2016 is the foundation for Microsoft's data strategy. With this new release, Microsoft delivers an end-to-end data management and business analytics solution, for demanding business applications as well as insights into business data on any device.

ReactJS

ReactJS is an open-source JavaScript library that provides a view for data rendered as HTML. React views are typically rendered using components that contain additional components specified as custom HTML tags.

Redis

Redis is an open source, in-memory data structure store which can be used as a database, cache or message broker. It supports data structures such as strings, hashes, lists, sets, sorted sets with range queries, bitmaps, hyperlog logs and geospatial indexes with radius queries. Redis has built-in replication, Lua scripting, LRU eviction, transactions, different levels of on-disk persistence and provides high availability via Redis Sentinel and automatic partitioning with Redis Cluster.

Scala

Scala is a popular object-oriented and functional programming language that has seen good adoption, especially for developing reactive applications. Also, Scala tends to be the programming language of choice for analytics applications that leverage the Apache Spark ecosystem.

Spring Boot

Spring Boot enables developers to build Spring applications quickly by taking a convention over configuration approach. The Spring Boot ecosystem is very vibrant and can be used to help realize a microservices-based architecture.

Spring Cloud

Spring Cloud is a collection of tools from Pivotal that provides solutions for quickly building some of the patterns commonly encountered in distributed systems. Spring Cloud builds upon some of the common building blocks of the Spring ecosystem and the Netflix Common Runtime Services and Libraries.

TypeScript

TypeScript is a language for application-scale JavaScript. It adds optional types, classes and modules to JavaScript and supports tools for large-scale JavaScript applications for any browser, host or OS. TypeScript compiles to readable, standards-based JavaScript.

WebRTC *(Web Real-Time Communication)*

WebRTC is a collection of communications protocols and application programming interfaces that enable real-time communication over peer-to-peer connections. This allows web browsers to not only request resources from backend servers, but also real-time information from browsers of other users. This enables construction of applications such as video conferencing, file transfer, chat or desktop sharing without the need of either internal or external plugins.

xUnit.net

xUnit.net is a free, open-source community-focused unit testing tool for the .NET Framework. Written by the original inventor of NUnit, xUnit.net is the latest technology for unit testing C#, F#, VB.NET and other .NET languages. xUnit.net works with ReSharper, CodeRush, TestDriven.NET and Xamarin. It is also a part of the .NET Foundation, and operates under their code of conduct.

Experiment

C#7.0

C# 7.0 adds a number of new features and brings a focus on data consumption, code simplification and performance. Few of the newly added features include: Tuples, Pattern Matching, Local Functions, Out Parameters and Literal Improvements, etc. Perhaps the most useful features with significant benefits would be Tuples and Pattern Matching.

Consul by Hashicorp

Consul is a versatile toolkit for discovering and configuring services within your infrastructure. It provides several key capabilities like service-discovery, health-checks, key-value store etc. Consul is designed to be friendly to both traditional application developers and the DevOps community, making it perfect for modern, elastic and highly available applications.

Go

Go is an expressive, concise, clean, and efficient general purpose programming language. It compiles quickly to machine code, yet has the convenience of garbage collection and the power of run-time reflection. It's a fast, statically typed, compiled language that feels like a dynamically typed, interpreted language. It has gained significant prominence in server software design and construction.

Java 9

Java 9 is one of the most awaited Java releases that will introduce several new features to an already rich programming ecosystem. The most significant of them are the introduction of jShell, Java Microbenchmarking Harness (JMH), HTTP 2.0 support, Java Platform Module System (JPMS) and several language and API improvements that should benefit the general Java developer.
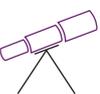
Microsoft Nano Server

Nano Server is a lightweight operating system, optimized for running "cloud-native" applications based on containers and micro-services. It can also be used as an agile and cost-effective data-center host with a dramatically smaller footprint, supporting remotely managed installation and optimizations for the cloud and DevOps workflows.

Microsoft Teams

Microsoft Teams is a new chat-based work-space in the Office 365 ecosystem. It is an entirely new experience that brings together people, conversations and content along with tools that teams need to easily collaborate and achieve more. The service is a cloud-based app that integrates well with familiar Office applications and is available to paying Office 365 customers. The Office 365 component adds a group chat tool to the company's office suite, supports both persistent and threaded chats to keep everyone engaged. It also acts as a hub for teamwork, which brings multiple good features of Office 365 together in one place. It is customizable at team level to enforce two-factor authentication, single sign on through Active Directory and encryption of data in transit or at rest.

Microsoft Visual Studio 2017

Enhancements to code navigation, IntelliSense, refactoring, code fixes, and debugging, saves time and effort on everyday tasks regardless of language or platform. For teams embracing DevOps, Visual Studio 2017 streamlines inner loop and speeds up code flow with brand new real time features such as live unit testing and real-time architectural dependency validation.

MS ASP.NET WebHooks

WebHooks is a lightweight HTTP pattern that provides a simple publish-subscribe model for wiring together Web APIs and SaaS services.

MS SharePoint and Office apps

The enhanced Microsoft Graph exposes multiple APIs from Microsoft cloud services to help businesses build smarter solutions. Improvements to the add-in model include support for MS Office on the Mac OS 2016, centralized deployment and programmatic creation of solution-specific ribbons, buttons and new extensibility features for Skype and Office 365 Groups.

Project Orleans

Project Orleans is a framework designed by Microsoft for building distributed high-scale computing applications using the actor model.

Vault by Hashicorp

Vault is a tool for securely accessing secrets. A secret is anything that you want to tightly control access to, such as API keys, passwords, certificates, and more. Vault provides a unified interface to any secret, while providing tight access control and recording a detailed audit log.

 # Watch

.Net Core 2.0

.NET Core is a general purpose, modular, cross-platform and open-source implementation of .NET. It includes a runtime, framework libraries, compilers and tools that support a variety of processor and OS targets. These components can be used together or separately. .Net Core 2.0 improves performance, .NET core tooling, lowering the barrier for cross platform dependency resolutions and has better support while running over Azure.

Kitura

Kitura is a new, modular, package-based web framework written in the Swift language. It is a high performance and simple to use web framework for building modern Swift applications.

Power BI

Microsoft Power BI is a collection of online services and features that enable you to find and visualize data, share discoveries and collaborate in intuitive new ways.

Swift on Server

Swift is a high-performance system programming language. It has a clean and modern syntax, offers seamless access to existing C and Objective-C code and frameworks, and is memory safe by default. Although inspired by Objective-C and many other languages, Swift is not itself a C-derived language. As a complete and independent language, Swift packages core features like flow control, data structures, and functions, with high-level constructs like objects, protocols, closures, and generics. Swift embraces modules, eliminating the need for headers and the code duplication they entail.

Universal Windows Platform

Windows 10 introduced the Universal Windows Platform (UWP), which evolves the Windows Runtime model and brings it into the Windows 10 unified core. As part of the core, the UWP now provides a common app platform that is available for every device that runs Windows 10. This package can be installed on a wide range of devices.
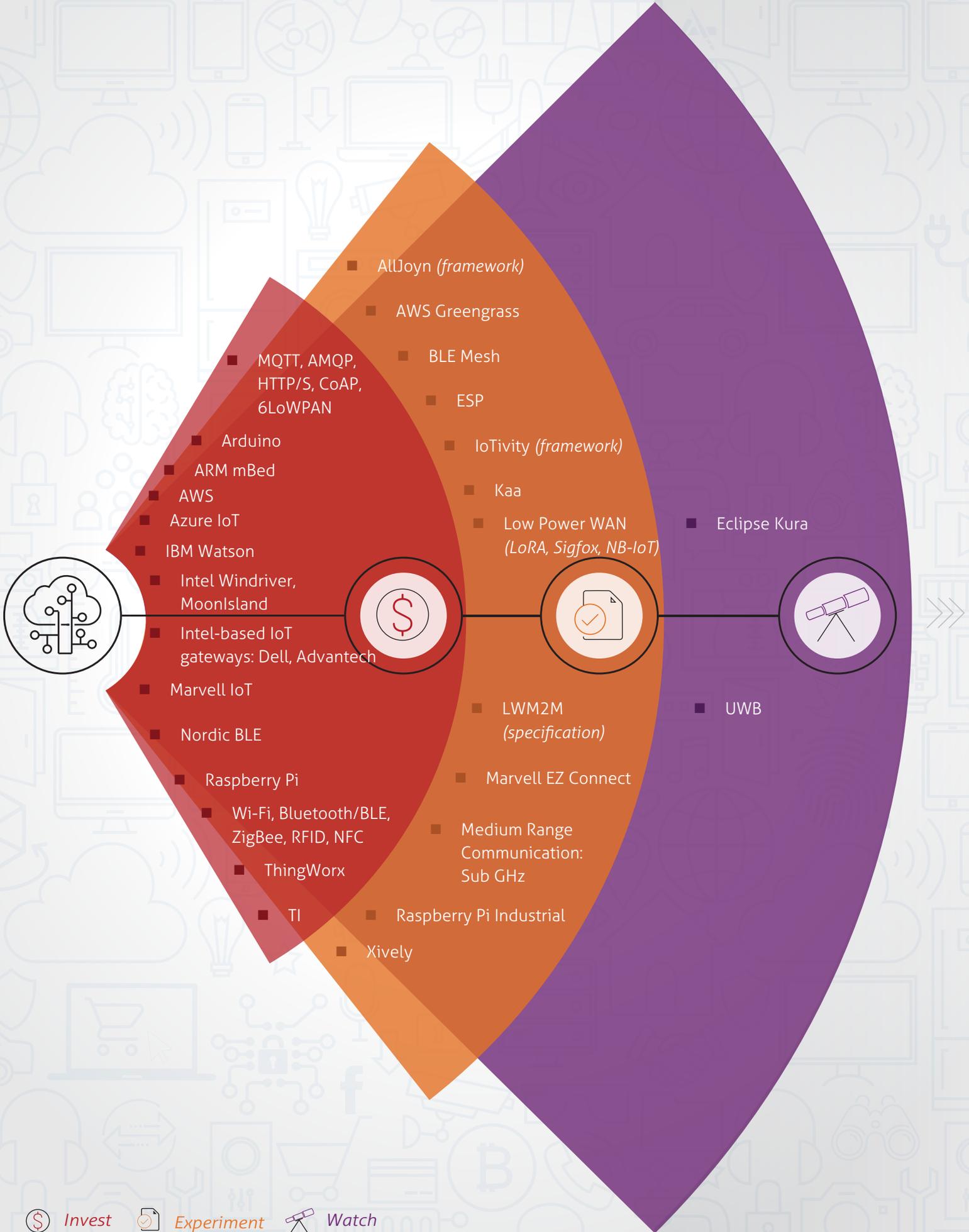
Windows Container

Windows Container is an isolated place where an application can run without affecting the rest of the system, and vice versa. Containers are the next evolution in virtualization. If you were inside a container, it would look very much like you were inside a freshly installed physical computer or a virtual machine. And, to Docker, a Windows container can be managed in the same way as any other container.

Windows Container Types:

- Windows Server Containers – provide application isolation through process and namespace isolation technology. A Windows Server container shares a kernel with the container host and all containers running on the host.

- Hyper-V Containers – expand on the isolation provided by Windows Server Containers by running each container in a highly optimized virtual machine. In this configuration the kernel of the container host is not shared with other Hyper-V Containers.

Internet of Things

AllJoyn (*framework*)

AWS Greengrass

BLE Mesh

MQTT, AMQP,
HTTP/S, CoAP,
6LoWPAN

ESP

Arduino

IoTivity (*framework*)

ARM mBed

Kaa

AWS

Low Power WAN
(*LoRA, Sigfox, NB-IoT*)

Eclipse Kura

Azure IoT

IBM Watson

Intel Windriver,
MoonIsland

Intel-based IoT
gateways: Dell, Advantech

Marvell IoT

LWM2M
(*specification*)

UWB

Nordic BLE

Raspberry Pi

Marvell EZ Connect

Wi-Fi, Bluetooth/BLE,
ZigBee, RFID, NFC

Medium Range
Communication:
Sub GHz

ThingWorx

TI

Raspberry Pi Industrial

Xively

$ *Invest*   *Experiment*   *Watch*

⑤ Invest

Application Protocols: MQTT, AMQP, HTTP/S, CoAP, 6LoWPAN

- **MQTT** – MQTT is a lightweight machine-to-machine communication protocol that uses a publish-subscribe messaging transport. Its small footprint is ideal for IoT devices.

- **AMQP** – Advanced Message Queuing Protocol (AMQP) is an open standard message-oriented middleware. It boasts of a low latency and high throughput queuing mechanism that is ideal for IoT workloads.

- **HTTP** – While HTTP is a rather heavy protocol for constrained devices, it continues to be used for IoT. REST over HTTP is a common technique for communication.

- **CoAP** – CoAP is a data transfer protocol designed for needs of constrained devices. It runs over UDP not on TCP and follows the client server model. CoAP is designed to inter-operate with HTTP and the RESTful web at large through simple proxies.

- **6LoWPAN** – 6LoWPAN is a low-power wireless mesh network where every node has its own IPv6 address, allowing it to connect directly to the Internet using open standards.

Arduino

Arduino is an open standards physical computing platform based on a simple microcontroller board and a development environment for writing software for the board. It is used in a variety of IoT use cases, such as taking inputs from sensors and switches, developing interactive objects and controlling a variety of subsystems. Arduino sketches enable quick development for prototype use cases. Where a Pi is suitable for prototyping gateways, Arduino is suitable for prototyping edge nodes.

ARM mBed

ARM mBed accelerates product development with a complete stack that includes an open standards embedded operating system, client and transport layer security, and an array of tools to develop and deploy software for ARM Cortex-M-based IoT devices. The overall ecosystem makes this an attractive choice.

AWS

The AWS IoT platform provides secure, bi-directional communication between Internet-connected things like the sensors, actuators, embedded devices or smart appliances, and the Amazon Web Services cloud. You can use it to collect, store, and analyze telemetry data from multiple devices, and create applications that enable users to control these devices from their phones or tablets. AWS' openness, richness of platform capability, and its robustness makes it a good IoT platform. Integration with other upcoming AWS technology such as Greengrass for edge, and Machine Learning, make it a good platform to build full IoT solutions.

Azure IoT

Microsoft Azure IoT Hub allows us to connect, monitor and manage millions of IoT assets from the Azure Cloud. With the addition of features like native MQTT v3.1.1 protocol support, X.509 certificate based device authentication and device management Azure IoT Hub continues to grow into a stable IoT platform. Combined with greater support for the open source stack, and big data capabilities around Spark and Hadoop, Azure becomes a good choice for building IoT solutions in general, and for building on the Microsoft stack in particular.

IBM Watson

IBM's PaaS offerings are seeing traction across industries. A major emphasis on transformational technologies like Cognitive, Blockchain and especially Watson IoT throw open a wide array of opportunities. Watson IoT's secure, standards-based and scalable software services both at the cloud and the edge can successfully address a very wide array of IoT use cases. A host of other services in the IBM Bluemix ecosystem can be easily integrated with Watson IoT, to deliver compelling solutions.

Intel Windriver, MoonIsland

MoonIsland is Intel's IoT Gateway Platform. It uses a series of Intel Quark Systems-on-Chip (SoC) based boards to connect to sensor networks. On an Intel ecosystem, these provide a good way to build a single-stack solution.

Intel-based IoT gateways: Dell, Advantech

In physical settings, gateways need to have both good computing power, and also be built robustly, to withstand conditions such as heat and dust. Dell and Advantech have a series of edge gateway products based on the Intel stack, which will be necessary to provide intelligence on the edge. These are based on Intel Quark and Atom processors. The equivalent prototype boards from Intel are Edison (based on Atom), and Galileo (based on Quark).

Marvell IoT

The Marvell 88MW300/302 Wi-Fi Microcontroller system-on-chip (SoC) with a full-featured micro controller built using ARM Cortex-M4F CPU and 802.11 b/g/n Wi-Fi. It has 512kB SRAM and an accelerated flash controller. It provides easy interfacing to sensors and actuators via a full set of I/O interfaces including SPI, I2C, UART, I2S, PWM, ADC and DAC. This chip enables smart connectivity with devices such as thermostats, air conditioners, appliances, lighting controls, mobile clients and other IoT devices. We believe the target for this will be towards IoT in smart homes.

Nordic BLE

Nordic Semiconductor provides a powerful, highly flexible multiprotocol SoC ideally suited for Bluetooth Smart and 2.4GHz ultra low-power wireless applications. We moved this from Experiment to Invest, based on robustness, and applicability to scenarios for indoor navigation, etc., that require a number of low-cost beacons to be created.

Raspberry Pi

Raspberry Pi is a low-cost, credit-card sized computer that plugs into a computer monitor or TV. It uses a standard keyboard and mouse, and is capable of doing everything you'd expect a desktop computer to do. It interfaces with USB, Ethernet, Bluetooth and Wi-Fi. The presence of SPI and GPIO banks allows the Pi to also be used to connect over serial, and to a variety of other sensors and actuators. We believe that the price point and flexibility make it a good prototyping platform.

Short Range Communication protocols/Standards: Wi-Fi, Bluetooth/BLE, ZigBee, RFID, NFC

Short range communication protocols are typically used in communication between sensor nodes and gateways. Based on scenario, different protocols and standards need to be used:

1. RFID/NFC: RFID and NFC technologies continue to see traction in areas like retail. Tags are being used extensively for low-cost identification purposes.

2. ZigBee: ZigBee is a low-cost, low-power, wireless mesh network standard targeted at the wide development of long battery life devices in wireless control and monitoring applications.

3. BLE: Bluetooth low energy (BLE), now marketed as Bluetooth Smart, is a wireless personal area network technology designed and marketed by the Bluetooth Special Interest Group. It targets the consumer space with novel applications for the healthcare, fitness, beacons, security and home entertainment industries. Bluetooth 5 now has mesh support which will help better serve IoT use cases.

ThingWorx

ThingWorx is a platform from PTC, well known in the world of machine-to-machine and IoT applications for the quick and intuitive development of IoT dashboards. ThingWorx accelerates IoT application development by compressing the design-develop-deploy cycles to reduce time to market and spur innovation.

TI

TM4C12x ARM® Cortex®-M4F core-based Microcontrollers: With connectivity features & cloud connectivity, TM4C12x MCUs integrate a variety of peripherals enabling IoT gateway and node applications. Availability of Launchpad series for different ARM core microcontrollers, Booster packs for different connectivity and sensor modules makes it easy to get started with development and prototyping. The chipset can subsequently be used for production fabrication.

Experiment

AllJoyn *(framework)*

AllJoyn is an open source software framework that makes it easy for devices and apps to discover and communicate with each other. Developers can write applications for interoperability regardless of transport layer, manufacturer, and without the need for Internet access. The wide industry backing, and its expansion beyond home automation makes it a good candidate to consider for standardization.

AWS Greengrass

AWS Greengrass is an edge software component that allows local compute, data filtering and secure communications between devices and the cloud. Developers can leverage the freedom to use familiar programming languages to easily execute program logic via AWS Lambda functions on the edge. While this service is available in limited preview, we believe the service has adequate merit to be included in our 'Experiment' category.

BLE Mesh

BLE Mesh is a key enabling technology for IoT, since this removes the single hop topology restrictions and enable the communication from every device to every other devices with in the mesh network. BLE mesh is based on existing low power BLE wireless technology so the devices can run on low power.

ESP

ESP8266 is a Low-power, highly-integrated Wi-Fi solution embedded with 8Mbit flash. ESP8266EX is among the most integrated Wi-Fi chips in the industry. Measuring just 5mm x 5mm, ESP8266EX requires minimal external circuitry and integrates a 32-bit Tensilica MCU, standard digital peripheral interfaces, antenna switches, RF balun, power amplifier, low noise receive amplifier, filters and power management modules - all in one small package. ESP32 is a 2.4 GHz Wi-Fi and Bluetooth combo chip TSMC low power 40nm technology. While initially targeted for prototyping, we have seen widespread adoption of this module, and so this features as "Experiment", with a potential to quickly move to "Invest."

IoTivity *(framework)*

IoTivity is an open-source software framework enabling seamless device-to-device connectivity to address the emerging needs of the Internet of Things. Backed by Samsung and Intel, it is designed to work on Wi-Fi Direct, BT/BTLE, ZigBee and ZWave. As AllJoyn and IoTivity are merged, current devices running either AllJoyn or IoTivity will be interoperable and backward compatible.

Kaa

Kaa is an open-source IoT Platform that promises to cater to a wide range of business domains. We found its ability to integrate with client provided custom code as one of its strong points. It also addresses core communication concerns and lets developers focus on the business logic around their IoT use cases. Its scalable server architecture is another strong point. Strong community support and noteworthy interest across industries makes this IoT platform an interesting area of experimentation.

Low Power WAN *(LoRA, Sigfox, NB-IoT)*

LPWAN technologies are intended for the use to connect low cost, low power and low bandwidth devices. Competing technologies continue to emerge in the LPWAN segment which uses Sub GHz frequency range, with the focus maximizing the range and minimizing the transmission power. Many companies are building business model around this. We believe this is an important component of long range communication, where solutions require non-cellular connectivity. Smart cities are another candidate for LPWAN solutions. LoRA and Sigfox are more established compared to NB-IoT. LoRA has seen greater adoption, and is a candidate that we believe will move quickly from "Experiment" to "Invest."

LWM2M *(specification)*

LWM2M is a new open industry protocol developed by OMA to provide a means to remotely perform service enablement and application management for embedded devices and connected appliances. It is a communication protocol for use between client software on an M2M device and server software on an M2M management and service platform. The backing by OMA, and its rich specification will drive its adoption to a large extent.

Marvell EZ-Connect

The Marvell EZ-Connect platform is helping to rapidly bring innovative devices to the fast-growing IoT market. Built on Marvell's silicon and coupled with industry standards certifications and leading Kinoma software, EZ-Connect is widely considered the most complete of industry-leading connectivity portfolios. It's hardware and software solutions are capable of enabling more efficient time-to-market IoT designs.

Medium-Range Communication: Sub GHz

Sub-GHz technology is an ideal choice for wireless applications requiring long range low power consumption and low data rate. This can transmit data to distant hubs, often few miles away, without hopping from node to node. This long-range transmission capability reduces the need for multiple expensive base stations or repeaters. This can be an alternative to LoRA, without requiring a tie-in into that ecosystem.

Raspberry Pi Industrial

The Compute Module is a Raspberry Pi in a more flexible form factor, intended for industrial application. CM3 is the Raspberry Pi 3-based revision of the original Compute Module. This comes with BCM2837 processor, 1GB RAM and a 4GB eMMC Flash device all integrated on to a small 67.6mm x 31mm board which fits into a standard DDR2

SODIMM connector. This is a good candidate to use when we are planning to productize our own hardware platform which is prototyped using Raspberry Pi3. This compares favorably with platforms like Advantech from a cost standpoint.

Xively

Xively is LogMeIn's Platform as a Service (PaaS) offering for IoT. Xively allows devices to securely connect at scale to transmit and store data.
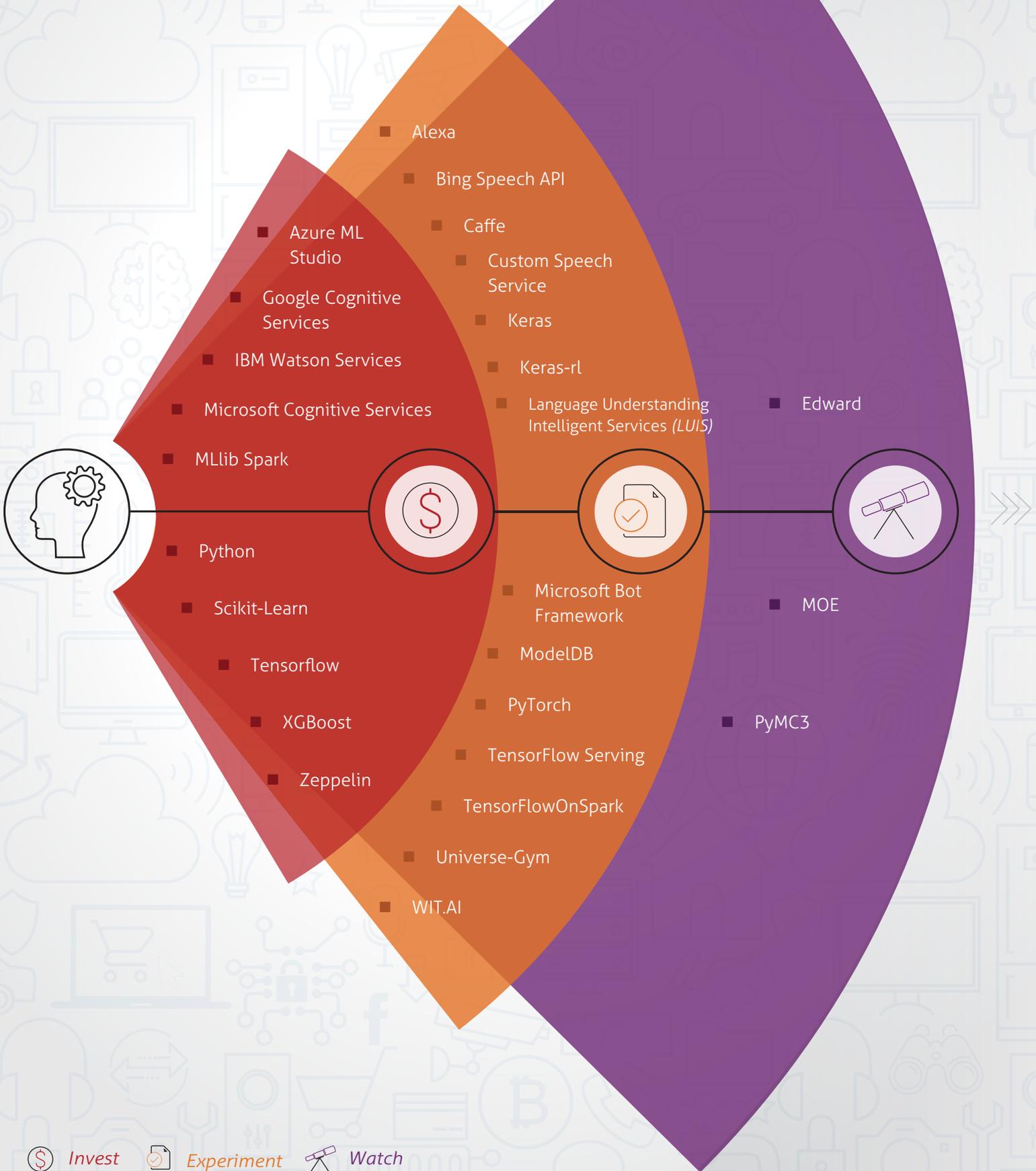

Watch

Eclipse Kura

Eclipse Kura is an Eclipse IoT project that provides a platform for building IoT gateways. It is a smart application container that enables remote management of such gateways and provides a wide range of APIs for allowing you to write and deploy your own IoT application. Kura runs on top of the Java Virtual Machine (JVM) and leverages OSGi, a dynamic component system for Java, to simplify the process of writing reusable software building blocks. Kura APIs offer easy access to the underlying hardware including serial ports, GPS, watchdog, USB, GPIOs, I2C, etc. It also offer OSGi bundle to simplify the management of network configurations, the communication with IoT servers, and the remote management of the gateway.

UWB

Specifically for indoor navigation, high bandwidth and extremely short pulses waveforms help in reducing the effect of multi-path interference and facilitate determination of ToA (Time of Arrival) for burst transmission between transmitter-receiver pairs. This makes UWB a good solution for indoor positioning. UWB signals are able to effectively measure distance between two devices up to 10-cm of accuracy, compared to roughly 5-m accuracy for Wi-Fi and Bluetooth. This needs to be looked at in relation to adoption (UWB by itself has not seen significant adoption), and power consumption, and so we place it in the "Watch" category.

Cognitive Computing

Azure ML Studio

Google Cognitive Services

IBM Watson Services

Microsoft Cognitive Services

MLlib Spark

Python

Scikit-Learn

Tensorflow

XGBoost

Zeppelin

Alexa

Bing Speech API

Caffe

Custom Speech Service

Keras

Keras-rl

Language Understanding Intelligent Services *(LUIS)*

Microsoft Bot Framework

ModelDB

PyTorch

TensorFlow Serving

TensorFlowOnSpark

Universe-Gym

WIT.AI

Edward

MOE

PyMC3

ⓢ Invest

Azure ML Studio

Azure ML allows users to create and train models, then turn them into APIs that can be consumed by other services

Google Cognitive Services

Google Cloud Vision API has powerful machine learning models to understand the content of an image. It exposes an easy to use REST API. It quickly classifies images into thousands of categories, detects individual objects and faces within images, and finds and reads printed words contained within images.

Google Cloud Speech API: Google Cloud Speech API enables developers to convert audio to text by applying powerful neural network models in an easy to use API.

The API recognizes over 80 languages and variants. It can help transcribe the text of users dictating to an application's microphone, enable command-and-control through voice, or transcribe audio files, among many other use cases.

Google Cloud Natural Language API: Google Cloud Natural Language API reveals the structure and meaning of text by offering powerful machine learning models in an easy to use REST API.

It can be used to extract information about people, places, events and much more, mentioned in text documents, news articles or blog posts.

It can be used to understand sentiment about products on social media or parse intent from customer conversations happening in a call center or a messaging app.

IBM Watson Services

Watson and its cognitive capabilities mirror some of the key cognitive elements of human expertise.

Watson provides a variety of smart services that help in building compelling cognitive applications. Watson has its strengths in natural language understanding including understanding context around language.

Watson provides services for conversation, dialog, document conversion, language translation, natural language classification, tone analysis etc.

Watson has Retrieve & Rank services which for example, can help an experienced technician to quickly find solutions from dense product manuals.

Watson AlchemyLanguage provides a set of APIs that offer text analysis through natural language processing. These services are exposed as easy to use REST APIs.

Microsoft Cognitive Services

Microsoft Cognitive services provide a suite of powerful machine learning APIs for computer vision, speech, language understanding and text analysis. These powerful services are exposed as easy to use REST interfaces. Computer vision APIs include image classification, recognizing celebrities, face and gender detection, detection of text in images and even classification of frames in real-time video. Microsoft offers services for content moderation including image, text and video. Microsoft speech services have very powerful speech services where the language model can be fine-tuned and customized by tailoring it to the vocabulary of the application and the speaking style of your users. It supports language detection, translation services and even custom translation system.

MLlib Spark

The goal of Apache Spark's machine learning library is to make practical machine learning scalable and easy. It consists of common learning algorithms and utilities, including classification, regression, clustering, collaborative filtering, dimensionality reduction, as well as lower-level optimization primitives and higher-level pipeline APIs.

Python

Python is the most preferred language by Data scientists today especially in the Big Data environments. The closer you get to working in an engineering environment, the more likely it is you might prefer Python. You can use Python when your data analysis tasks need to be integrated with web apps or if statistics code needs to be incorporated into a production database. It is a common ask to have Python knowledge when we work in distributed environments.

Scikit-Learn

Scikit-Learn is a well-designed, open-source tool for data analysis, data mining and machine learning. It is completely Python based and leverages Python packages like Numpy and Scipy. The library has a consistent programming interface and easy to get started.

Tensorflow

Tensorflow is a popular, general, flexible, open-source framework for numerical computation using data flow graphs that can be deployed on CPUs and GPUs. It can used to build various types of neural network architectures for deep learning. It is written in C++ and has Python bindings. The flexible architecture allows the deployment of computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API. TensorFlow supports multi-GPU training and the servers can be grouped into a cluster with the ability to distribute a computation task across the cluster. It integrates with a suite of visualization tool called TensorBoard which allows for better understanding, optimizing and quantitative metric plotting of data being trained by neural networks.

XGBoost

XGBoost is a machine library using gradient boosted decision trees designed for speed and performance. It supports regression, classification, ranking and user-defined objectives. It is scalable, supports parallel and distributed execution and provides interfaces to multiple programming languages. XGBoost is very effective for prediction tasks on tabular or structured data-sets. XGBoost library is used in many data science and machine learning challenges.

Zeppelin

Zeppelin is a modern web-based tool for the data scientists to collaborate over large-scale data exploration and visualization projects. It is a notebook style interpreter that enables collaborative analysis sessions sharing between users and is independent of the execution framework itself. Current version runs on top of Apache Spark, but it has a pluggable interpreter APIs to support other data processing systems.

Experiment

Alexa

Alexa is Amazon's voice control system with a built-in natural-language processing system. Alexa is built in the cloud, so it is always getting smarter. The more customers use Alexa, the more she adapts to speech patterns, vocabulary and personal preferences.

Bing Speech API

It is a cloud-based API that provides advanced algorithms to process spoken language. With Bing Speech API, you can add speech-driven actions to your apps, including real-time interaction with the user. Bing Speech APIs and libraries enable speech capabilities on all internet-connected devices. Every major platform including Android, iOS, Windows, and 3rd-party IoT devices are supported. It offers industry-leading speech-to-text, text-to-speech, and language understanding capabilities delivered through the cloud. Microsoft uses Bing Speech API for Windows applications like Cortana and Skype Translator as well as Android applications like Bing Torque for Android Wear and Android Phone.

Caffe

Caffe is an open-source Deep Learning framework from Berkeley Vision Learning Center. It has good support for Convolutional Neural Networks which is commonly used for vision tasks in deep learning. Models and optimization in Caffe are defined by configuration rather than by code. It has a growing library of pre-trained models and provides Python, Matlab and C++ interfaces.

Custom Speech Service

Custom Speech Service enables you to create customized language models and acoustic models tailored to your application and your users. By uploading your specific speech and/or text data to Custom Speech Service, you can create custom models that can be used in conjunction with Microsoft's existing state-of-the-art speech models.

Keras

Keras is a high-level neural networks API, written in Python and capable of running on top of either TensorFlow or Theano. It is open-source and has a modular design which can be easily extended with custom modules. It has a high-level interface that enables rapid prototyping of complex neural network architectures. Keras supports convolutional networks, recurrent networks and a combinations of the two.

Keras-rl

Keras-rl implements state-of-the-art deep reinforcement learning algorithms in Python and seamlessly integrates with the deep learning library Keras. Keras-rl makes it really easy to run state-of-the-art deep reinforcement learning algorithms, uses Keras and thus supports Theano or TensorFlow backend. Hence algorithms can efficiently be trained either on CPU or GPU. It works with OpenAI Gym out of the box as well and makes evaluating and playing around with different algorithms relatively easy. Keras-rl is extendable which makes it easy to implement your own environments and even algorithms by simply extending some simple abstract classes.

Language Understanding Intelligent Services *(LUIS)*

LUIS is designed to enable developers to build smart applications that can understand human language and accordingly react to user requests. With LUIS, a developer can quickly deploy an HTTP endpoint that will take the sentences sent to it and interpret them in terms of their intents (the intentions they convey) and entities (key information relevant to the intent). LUIS supports several languages as English, French, Italian, German, Spanish, Brazilian, Portuguese, Japanese, Korean and Chinese. Many of the common chat bots can be easily built using LUIS.

Microsoft Bot Framework

The Microsoft Bot Framework provides just what you need to build and connect intelligent bots that interact naturally wherever your users are talking, from text/sms to Skype, Slack, Office 365 mail and other popular services. The Framework provides developers with a developer portal and SDK to build Bots, a Bot Connector service to connect to social channels such as Twitter and Slack, and a Bot Directory to discover and use existing bots.

ModelDB

ModelDB is a machine learning model management tool. Machine learning is an iterative process and often 100s of models are built and evaluated before arriving at a model that meets the acceptance criteria. ModelDB is an end-to-end system that tracks models as they are built, extracts and stores relevant metadata (e.g., hyperparameters, data sources) for models, handles versioning and makes this data available for easy querying and visualization. ModelDB integrates with popular machine learning libraries like scikit-learn and spark-ml.

PyTorch

PyTorch is a Python-based tensor computing framework with GPU accelaration. PyTorch is a fast, minimalist network which is easy to extend with custom neural network modules written in Python. Compared to other deep learning frameworks, PyTorch is more easier to debug and understanding error messages and stack traces is comparatively easier. PyTorch supports building dynamic neural networks which is built on tape-based autograd system.

TensorFlow Serving

TensorFlow Serving makes the process of taking a machine learning model into production easier and faster. It allows safe deployment of new models and running experiments while keeping the same server architecture and APIs. It enables model lifecycle management, facilitates experimentation and adds versioning support. Out of the box, it provides integration with TensorFlow, but it can be extended to serve other types of models.

TensorFlowOnSpark

TensorFlowOnSpark enables merging of deep learning with big data platforms.

TensorFlowOnSpark enables distributed TensorFlow training and inference on Apache Spark clusters. TensorFlowOnSpark enables distributed deep learning on a cluster of GPU and CPU servers. It was developed at Yahoo and then contributed to the open-source community. It seeks to minimize the amount of code changes required to run existing TensorFlow programs on a shared grid. Its Spark-compatible API helps manage the TensorFlow cluster.

Universe-Gym

Universe allows an AI agent to use a computer like a human does: by looking at screen pixels and operating a virtual keyboard and mouse. Universe lets you train a single agent on any task a human can complete with a computer. Universe exposes a wide range of environments through a common interface: the agent operates a remote desktop by observing pixels of a screen and producing keyboard and mouse commands. Existing game environments and real-world browser tasks can be controlled from Universe-GYM environment. Gym is a toolkit for developing and comparing reinforcement learning algorithms. With Universe, any program can be run within a Gym environment.

WIT.AI

Wit.ai makes it easy for developers to build applications and devices that you can talk or text to. It empowers developers with an open and extensible natural language platform. Wit.ai learns human language from every interaction, and leverages the community: what's learned is shared across developers. Wit.ai is free for personal and commercial usage.
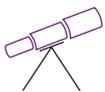
Watch

Edward

Edward is a Python library for probabilistic modeling, inference, and criticism. Edward is built on TensorFlow and is a testbed for fast experimentation and research. It supports probabilistic models, ranging from classical hierarchical models on small data sets to complex deep probabilistic models on large data sets. Edward fuses Bayesian statistics and machine learning, deep learning, and probabilistic programming. It also supports modeling with Directed Graphical models, Neural Networks, intractable likelihood, probabilistic programs and Bayesian non-parametrics. Edward supports multiple mechanisms for inference including variational inference, inference using Monte Carlo techniques and has techniques for criticism of model and inference.

MOE

Evaluating hyper-parameters for machine learning models is time-consuming. MOE helps tune model hyper-parameters faster and more accurate. MOE uses Bayesian techniques for learning optimal parameters. It is open-source and exposes Python and REST interfaces.
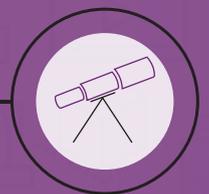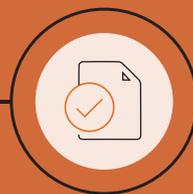
PyMC3

Probabilistic Programming (PP) allows for specification of Bayesian statistical models in code and Bayesian inference over those models. PyMC3 is an open-source, Python package and a PP framework with an intuitive, readable yet powerful syntax that is close to the natural syntax statisticians use to describe models. PyMC3 supports modeling, inference and features next-generation Markov chain Monte Carlo (MCMC) sampling algorithms that work well in high-dimensional data. It also support variational inference and is an excellent environment for developing fully Bayesian Gaussian Process models.

Blockchain

Blockchain as
a Service *(BaaS)*

Eris by MONAX

Enterprise Ethereum
Alliance

Fabric by
Hyperledger

Quorum

Hyperledger by R3

$ Invest   Experiment   Watch

Experiment

Blockchain as a Service *(BaaS)*

Blockchain as a Service (BaaS) provides a rapid, low-cost, low-risk and fail-fast platform for organizations to collaborate together by experimenting with new business processes—backed by a cloud platform with the largest compliance portfolio in the industry.

Eris by MONAX

Eris is a platform for building, testing, maintaining and operating ecosystem applications with a Blockchain backend. Eris makes it easy and simple to wrangle the dragons of smart contract Blockchains.

Fabric by Hyperledger

Hyperledger Fabric is a platform for distributed ledger solutions, underpinned by a modular architecture delivering high degrees of confidentiality, resiliency, flexibility and scalability. It is designed to support pluggable implementations of different components and accomodate the complexity and intricacies that exist across the economic ecosystem.

Quorum

Quorum is ideal for any application requiring high-speed and high-throughput processing of private transactions within a permissioned group of known participants. Quorum addresses specific challenges to Blockchain technology adoption within the financial industry, and beyond.


Watch

Enterprise Ethereum Alliance

The Enterprise Ethereum Alliance connects Fortune 500 enterprises, startups, academics, and technology vendors with Ethereum subject-matter experts. This alliance is looking to build a private implementation derived from the open-source public implementation of the Ethereum code base.

Hyperledger by R3

The Hyperledger project is an open-source collaborative effort created to advance cross-industry Blockchain technologies. It is a global collaboration including leaders in finance, banking, IoT, supply chain, manufacturing and technology.

VR/AR
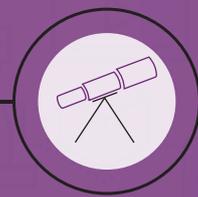
Blender

Daydream

Google VR

Google Tango

.HoloLens

WebVR

Unity

Vuforia

Experiment

Blender

Blender is a free and open-source tool to create 3D models which can be used to create animations, visual effects, art, 3D printed models, interactive 3D applications as well as video games. It is available for Windows, Linux and MacOS.

Daydream

The Daydream View as well as Daydream controller was launched alongside with Google Pixel.

Contrasting from Google's first VR platform, Google Cardboard, Daydream is built into the Android operating system starting with the release of Android 7.1 Nougat. The platform includes both software and hardware specifications, designating compatible phones "Daydream-ready". Daydream will only work on certain newer phones with specific components.

Daydream headsets are packaged with a wireless controller. This controller can be used for interacting with the virtual world through button presses or through waving the device. On-board sensors are used to track the orientation of the controller and approximate the position of the user's hand.

Google VR

Formerly known as Google Cardboard, known to be the cheapest VR headwear. Users can either build their own viewer from simple, low-cost components using specifications published by Google, or purchase a pre-manufactured ones. To use the platform, users run Cardboard-compatible applications on their phones, place the phone into the back of the viewer, and view content through the lenses.

HoloLens

HoloLens is a true Mixed Reality smart-glasses developed by Microsoft. It gained popularity for being one of the first computers running the Windows Holographic platform under the Windows 10 operating system. The device has a lot of potential to be helpful in vast domains like real estate, architecture, education, enterprise/industrial training, defense, gaming, travel and transportation, etc. This supports both 2D as well as 3D applications. The 2D applications run in a slate, and 3D with complete immersive experience. HoloLens provides GGV mode of interaction – Gaze, Gesture and Voice. MS Visual Studio is used to develop an app for this device, along with few more software like Unity and Vuforia.

Unity

Unity is a cross-platform game engine developed by Unity Technologies which is used for creating 2D/3D applications, video games, animations for websites, gaming consoles, mobile devices, etc. Unity is notable for its ability to target games to multiple platforms. Within a project, developers have control over delivery to mobile devices, web browsers, desktops, and consoles. This tool is also used for creating content for VR/AR devices – Google VR, HoloLens, etc.

Vuforia

Vuforia is an Augmented Reality SDK for devices which enables the creation of Augmented Reality applications. It uses Computer Vision technology to recognize and track Image Targets and simple 3D objects, such as boxes, in real time. This image registration capability enables developers to position and orient virtual objects, such as 3D models and other media, in relation to real-world images when these are viewed through the camera of a mobile device/HoloLens.

The virtual object then tracks the position and orientation of the image in real time so that the viewer's perspective on the object corresponds with their perspective on the Image Target, so that it appears that the virtual object is a part of the real-world scene.
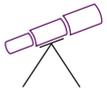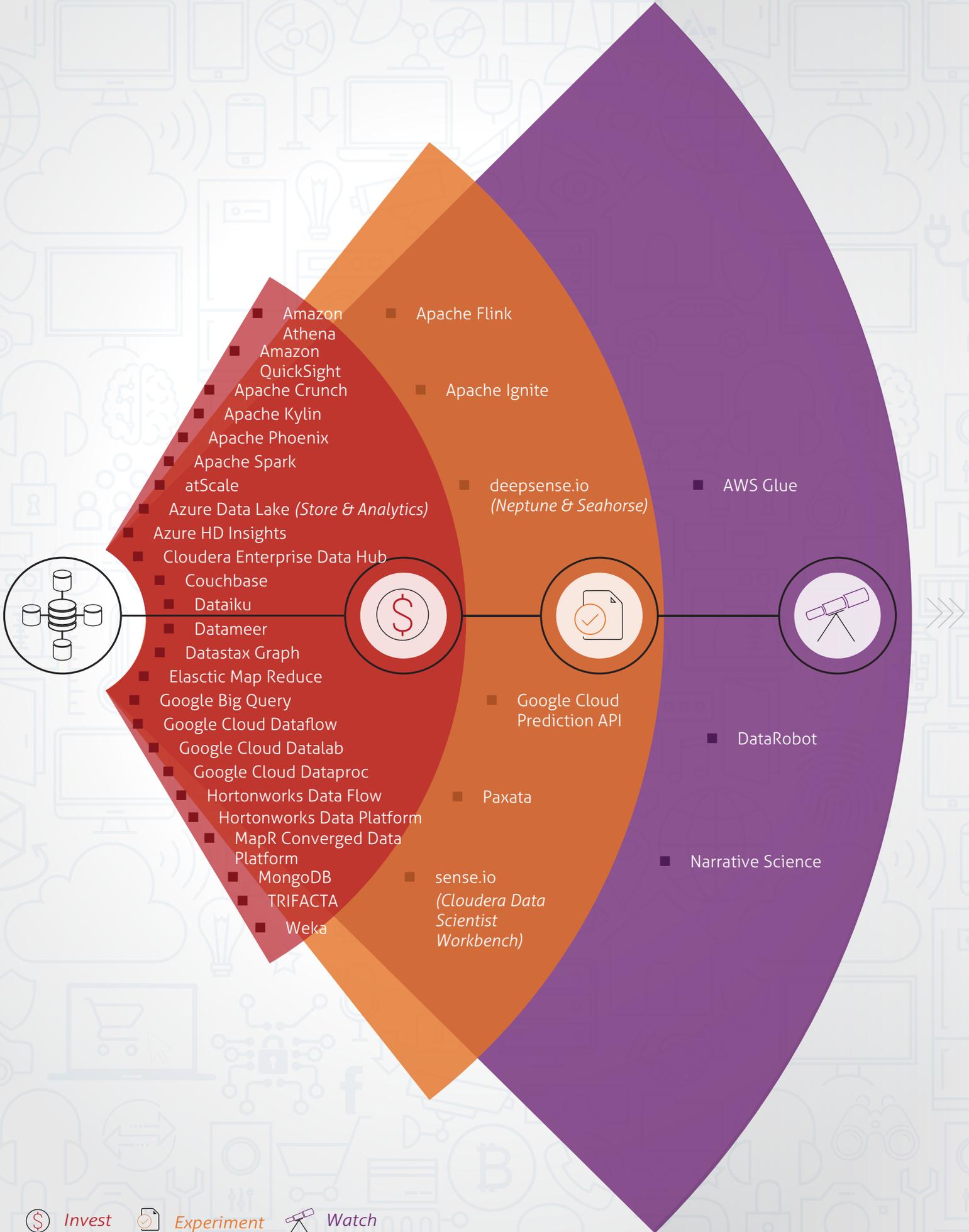
Watch

Google Tango

Tango is a project from Google that enable smartphones to create an augmented reality environment around the user. It uses computer vision to enable smartphones or tablets to detect their position in the augmented world with respect to real world. This allows the application developers to create great experiences like indoor navigation, distance measurement, mapping the environment, etc. The first two of the commercial devices are Lenovo Phab 2 Pro and Asus Zenfone AR. The Zenfone AR also has capability to support Daydream view kit, which makes it even more advanced among the typical VR-enabled devices.

WebVR

WebVR is an experimental JavaScript API to enable VR experience over browsers. Currently available as a beta release for major browsers, this eliminates the installation of additional apps which could be memory intensive. The WebVR API exposes certain interfaces which allows the web applications to present their content in virtual reality by utilizing the camera settings and few more sensor data along with device integrations using WebGL.

Big Data & Analytics

Amazon Athena
Amazon QuickSight
Apache Crunch
Apache Kylin
Apache Phoenix
Apache Spark
atScale
Azure Data Lake *(Store & Analytics)*
Azure HD Insights
Cloudera Enterprise Data Hub
Couchbase
Dataiku
Datameer
Datastax Graph
Elasctic Map Reduce
Google Big Query
Google Cloud Dataflow
Google Cloud Datalab
Google Cloud Dataproc
Hortonworks Data Flow
Hortonworks Data Platform
MapR Converged Data Platform
MongoDB
TRIFACTA
Weka

Apache Flink

Apache Ignite

deepsense.io
(Neptune & Seahorse)

Google Cloud
Prediction API

Paxata

sense.io
(Cloudera Data Scientist Workbench)

AWS Glue

DataRobot

Narrative Science

⑤ Invest

Amazon Athena

Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to manage, and you pay only for the queries that you run. Athena is easy to use. Simply point to your data in Amazon S3, define the schema, and start querying using standard SQL. Most results are delivered within seconds. With Athena, there is no need for complex ETL jobs to prepare your data for analysis. This makes it easy for anyone with SQL skills to quickly analyze large-scale datasets.

Amazon QuickSight

Amazon QuickSight is a fast, cloud-powered business analytics service that makes it easy to build visualizations, perform ad-hoc analysis and quickly get business insights from your data. Using this cloud-based service, you can easily connect to your data, perform advanced analysis and create stunning visualizations with rich dashboards that can be accessed from any browser or mobile device.

Apache Crunch

The Apache Crunch Java library provides a framework for writing, testing and running MapReduce pipelines. Its goal is to make pipelines that are composed of many user-defined functions that are simple to write, easy to test and efficient to run.

Apache Kylin

Apache Kylin™ is an open-source Distributed Analytics Engine designed to provide SQL interface and multi-dimensional analysis (OLAP) on Hadoop supporting extremely large datasets.

Apache Phoenix

Apache Phoenix enables OLTP and operational analytics for Apache Hadoop by providing a relational database layer leveraging Apache HBase as its backing store. It includes integration with Apache Spark, Pig, Flume, Map Reduce and other products in the Hadoop ecosystem. It is accessed via a JDBC driver and enables querying, updating and managing HBase tables through standard SQL.

Apache Spark

Apache Spark is a fast and general engine for large-scale data processing. It offers high-level APIs in Java, Scala and Python as well as a rich set of libraries including stream processing, machine learning, and graph analytics.

atScale

atScale lets you scale out business intelligence (BI) on Big Data and OLAP on Hadoop with Virtual Cubes™. This enables you to create models with measures and dimensions just like OLAP on Big Data.

Azure Data Lake *(Store & Analytics)*

Data Lake Analytics let you easily develop and run massively parallel data transformation and processing programmes in U-SQL, R, Python and .NET. With no infrastructure to manage, you can process data on demand and scale instantly while only paying per job.

Azure HD Insights

HDInsight is the only fully-managed Hadoop offering in the cloud that provides optimized open-source analytic clusters for Spark, Hive, MapReduce, HBase, Storm, Kafka and R Server backed by a 99.9% SLA.

Cloudera Enterprise Data Hub

Cloudera Enterprise, powered by Apache Hadoop and Apache Spark at the core, enables an enterprise data hub providing security, governance, management, support, and commercial ecosystem required for production success. Importantly a strong ecosystem of partners and in-house tool makes it a powerful market player and influencer.

Couchbase

Couchbase is an open-source, document-oriented NoSQL database for modern web, mobile and IoT applications. It is designed for ease of development and Internet-scale performance.

Dataiku

Dataiku is a collaborative data science platform to turn raw data into predictions.

Datameer

Datameer's end-to-end big data analytics platform for Hadoop empowers businesses to directly integrate, analyze and visualize data.

Datastax Graph

Datastax Graph is a scalable, real-time graph database built around Datastax Cassandra. It is one-of-a-kind distributed graph process applicable to variety of use cases such as customer 360, customer journey social maps and so on. Datastax Enterprise Graph is a multi-model platform, which provides support for key-value, tabular, JSON/Document and graph data models. It also incorporates all of the enterprise-class functionality such as advanced security, built-in analytics, enterprise search functionality, visual management, monitoring and development tools.

Elasctic Map Reduce

Amazon EMR provides a managed Hadoop framework that makes it easy, fast and cost-effective to process vast amounts of data across dynamically scalable Amazon EC2 instances. One can also run other popular distributed frameworks such as Apache Spark, HBase, Presto and Flink in Amazon EMR and interact with data in other AWS data stores.

Google Big Query

Big Query is a fast, economical and fully-managed enterprise data warehouse for large-scale data analytics

Google Cloud Dataflow

Google Cloud Dataflow is a unified programming model and a managed service for developing and executing a wide range of data processing patterns including ETL, batch computation and continuous computation. Cloud Dataflow frees you from operational tasks like resource management and performance optimization.

Google Cloud Datalab

Google Cloud Datalab is a powerful interactive tool created to explore, analyze, transform, visualize data and build machine learning models on the Google Cloud Platform.

Google Cloud Dataproc

Google Cloud Dataproc is a service for Apache Hadoop, Apache Spark, Apache Pig and Apache Hive to easily process big datasets at low cost.

Hortonworks Data Flow

Hortonworks Data Flow makes streaming analytics faster and easier, by enabling accelerated data collection, curation, analysis and delivery in real time. It is available for roll out on-premises or in the cloud through an integrated solution with Apache NiFi, Kafka and Storm.

Hortonworks Data Platform

Horton Data Platform is the industry's only true secure, enterprise-ready open-source Apache™ Hadoop® distribution based on a centralized architecture (YARN).

MapR Converged Data Platform

MapR Converged Data Platform integrates Hadoop, Spark and Apache Drill with real-time database capabilities, global event streaming and scalable enterprise storage to power a new generation of big data applications. The MapR Platform delivers enterprise grade security, reliability and real-time performance while dramatically lowering both hardware and operational costs of your most important applications and data.

MongoDB

MongoDB is a free and open-source cross-platform document-oriented database solution. Classified as a NoSQL database program, MongoDB uses JSON-like documents with schemas to manage large volumes of data.

TRIFACTA

Trifacta is a platform for preparing data for analysis and is compliant with the Hadoop ecosystem using Spark or MapReduce.

Weka

Weka is a data mining software written in Java, developed at the Machine Learning Group at University of Waikato, New Zealand. It is a GUI-based tool which is very good for beginners in data science and the best part is that it is open-source.

Experiment

Apache Flink

Apache Flink is an open-source system for expressive, declarative, fast and efficient data analysis. It combines the scalability and programming flexibility of distributed MapReduce like platforms with the efficiency, out-of-core execution and query optimization capabilities found in parallel databases.

Apache Ignite

The Apache Ignite in-memory Data Fabric is designed to deliver uncompromising performance for a wide set of in-memory computing use cases from high-performance computing, to the industry's most advanced data grid, in-memory SQL, in-memory file system, streaming and more.

deepsense.io *(Neptune & Seahorse)*

deepsense.io provides accelerated data sciences by automating workflows, job generation (Spark) and monitoring.

Google Cloud Prediction API

Google Cloud Prediction API offers RESTful APIs for building machine learning models for Android applications. This platform is specifically for mobile applications based on the Android OS.

Paxata

Paxata is one of the few organizations which focus on data cleaning and preparation. It however, does NOT support machine learning or statistical modeling. It is a MS Excel-like application that with visual guidance, makes it easy to bring together data, find/fix dirty or missing data, share and re-use data projects across teams. Like others mentioned here, Paxata eliminates coding or scripting, thereby overcoming technical barriers involved in handling data.

sense.io *(Cloudera Data Scientist Workbench)*

Cloudera Data-Scientist Workbench is a way to accelerate data science from exploration to production using R, Python, Spark and more. It helps in deploying pipelines and models on premise or in the cloud.
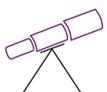
Watch

AWS Glue

AWS Glue is a fully-managed ETL service that makes it easy to move data between your data stores. AWS Glue simplifies and automates the difficult and time-consuming data discovery, conversion, mapping and job scheduling tasks. AWS Glue guides you through the process of moving your data with an easy-to-use console that helps you understand your data sources, prepare the data for analytics and load it reliably from data sources to destinations.

DataRobot

DataRobot (DR) is a highly-automated machine learning platform.

Narrative Science

Narrative Science is based on a unique idea in the sense that it generates automated reports using data. It works like a data story-telling tool which uses advanced natural language processing to create reports which are similar to a consulting report.

About Mindtree

Mindtree (NSE: MINDTREE) delivers technology services and accelerates growth for Global 1000 companies by solving complex business challenges with breakthrough technical innovations. Mindtree specializes in e-commerce, mobility, cloud enablement, digital transformation, business intelligence, data analytics, testing, infrastructure, EAI and ERP solutions. We are among the fastest growing technology firms globally with more than 200 clients and offices in 14 countries.