Para ver estadísticas de Internet más dinámicas, busque “estadísticas en vivo de Internet”.

<https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>

El sitio de IoT World Today y su lista de correo electrónico son un medio para recibir noticias y desarrollos relacionados con IoT.

https://www.credly.com/badges/7c9e31fe-1af8-4661-85d9-b63753cf906e/public\_url

Introducción al Internet de las Cosas y Transformación Digital

¿Sabía que los agricultores pueden colocar sensores en sus cultivos que les indiquen cuándo regarlos, cuánta agua es necesaria y cuándo cosechar? Con esta información, los agricultores pueden obtener la mejor calidad y cantidad de sus cultivos. Los mineros de carbón pueden colocar sensores en una mina que detecten pequeñas cantidades de gases peligrosos. Esta información salva vidas. Las compañías de seguros de automóviles pueden ofrecer a los conductores tarifas más bajas a cambio de acceso a sus datos de manejo. Esto permite una asignación de precios más justa y más precisa, y aumenta las ganancias al tiempo que reduce los costos.

IoT trata sobre datos. IoT trata sobre la digitalización de aspectos de nuestras vidas, nuestros negocios y nuestros gobiernos para proporcionar información procesable sobre cómo se pueden salvar vidas, cómo se pueden crear eficiencias y cómo se pueden mejorar las comunidades

* + 1. **La Evolución de la Transformación Digital**

En nuestro mundo actual, hay más dispositivos inteligentes que personas. Una cantidad cada vez más grande de personas está conectada a Internet de una forma o de otra, las 24 horas del día. Una cantidad cada vez mayor de personas posee y depende de tres, cuatro o más dispositivos inteligentes. Estos pueden incluir teléfonos inteligentes, monitores de ejercicio y salud, lectores electrónicos y tabletas. Se prevé que, en promedio, habrá 3,4 dispositivos inteligentes o conexiones para cada persona en la tierra.

IoT es relevante para muchos sectores. Los sistemas de IoT contribuyen a los controles ambientales, el comercio minorista, el transporte, los servicios de salud y las industrias agrícolas, entre muchos otros. Según Statista, se prevé que la cantidad de dispositivos de IoT en uso en todos los sectores relevantes aumente a más de ocho mil millones para 2030. En cuanto a los consumidores, las áreas importantes de crecimiento son Internet y los dispositivos de medios digitales, que incluyen los teléfonos inteligentes. Esta área también se prevé que crezca a más de ocho mil millones para 2030. Otras aplicaciones con más de un millón de dispositivos conectados son vehículos conectados y autónomos, infraestructura de TI, administración de activos y red eléctrica de servicios públicos.

**1.1.3 El impacto de la Transformación Digital en los Negocios**

**Automatización del hogar inteligente**

La tecnología digital ha permitido que las empresas innoven su enfoque para interactuar con la sociedad. Personas de todas las generaciones se sienten más cómodas con la tecnología digital y utilizan dispositivos inteligentes para su beneficio en los días atareados.

**Edificios inteligentes**

Muchas empresas ahora proporcionan algunos o todos sus servicios en línea. Desde la comodidad de su hogar, el automóvil, el gimnasio o una oficina, puede comprar alimentos en línea, pedir comida de restaurante para que se entregue en la puerta de su casa, reservar viajes en línea, realizar pedidos de ropa, equipo de campamento, taxis, y mantenerse conectado con sus amigos o conocer a un nuevo novio.

Los sensores están en todas partes y generan grandes cantidades de datos.

Los hogares inteligentes pueden contar con sensores de movimiento, sensores de agua, sensores de luz, sensores de timbre y sensores de temperatura. Puede haber sensores en los semáforos, los camiones de transporte, los garages de estacionamiento, las cámaras de seguridad, los trenes y los aviones. Todos estos sensores y dispositivos de medición recopilan y transmiten sus propios datos. Los datos pueden almacenarse y analizarse en una fecha posterior o se pueden analizar inmediatamente para ser utilizados para modificar computadoras, dispositivos móviles o procesos de cualquier tipo.

¿Cómo se utilizan los datos almacenados y analizados?

Laboral

Las empresas determinan los patrones de compra, pronostican nuevas tendencias y optimizan la producción.

Gobierno

Los gobiernos monitorean el medio ambiente, pronostican las tendencias de la población, pronostican las tasas de criminalidad y planifican los servicios sociales.

Ciudad

Las ciudades controlan el tráfico, monitorean el estacionamiento, brindan apoyo policial o de bomberos más rápido y controlan la gestión de desechos.

**1.1.4 ¿Pueden pensar los dispositivos inteligentes?**

Todos los dispositivos digitales funcionan según los programas informáticos y los datos suministrados. La inteligencia artificial implica que estos dispositivos son capaces de pensar por sí mismos. Si se programan de manera correcta, los dispositivos inteligentes pueden evaluar los datos que reciben, y modificar los procesos o la configuración “en el acto”. Si se les proporcionan los datos suficientes, pueden “aprender” y modificar sus propios códigos según los nuevos parámetros.

Imagine un camión de transporte refrigerado que lleva productos congelados y está equipado con un sensor de posicionamiento global. Cuando el camión ingresa en una ciudad importante, el sensor determina que hay un accidente adelante que produce una congestión de tráfico grande. El sensor envía los datos al sistema informático que recopila los datos y toma las decisiones. Luego, el sistema alerta al conductor de las nuevas condiciones de modo que pueda evitarse el accidente. Esta interacción automática ahorró al conductor tiempo y hará que los productos transportados lleguen al mercado con mayor rapidez con un producto que siga estando congelado.

Las oficinas corporativas pueden estar ocupadas por miles de empleados. Mantener el entorno, como la iluminación, el calor, la humedad, en el edificio dentro de parámetros aceptables ayuda a mantener a los empleados felices y, por lo tanto, más productivos.

Las ciudades inteligentes, como Barcelona, en España, usan sensores para controlar muchos de sus sistemas de infraestructura, como el flujo de tráfico, el estacionamiento, la utilización del agua y los sistemas hídricos.

Ejemplo 1: sensores de peso en espacios de estacionamiento permiten que los conductores sepan con rapidez dónde hay un espacio disponible. Esto reduce el tiempo de conducción y ralentí para el conductor y disminuye las emisiones de carbono al medioambiente.

Ejemplo 2: sensores de semáforos pueden detectar la congestión de tráfico. Estos datos se envían al sistema informático que recopila los datos y toma las decisiones. Las nuevas decisiones se envían a los actuadores, que modifican los tiempos de iluminación de la luz roja frente a la luz verde para brindar asistencia en el flujo de tráfico. Esto no solo disminuye el tiempo de inactividad, sino que también reduce la frustración del conductor y los accidentes.

Los automóviles automáticos están revolucionando el transporte. Los automóviles están equipados con muchos sensores de ultrasonido, cámaras, GPS de precisión y computadoras. La combinación del equipo incorporado permite a las computadoras identificar otros automóviles, carriles, peatones y obstrucciones. Esta información permite que el automóvil permanezca en su carril, se detenga cuando sea necesario y sortee obstáculos.

Algunos de los fabricantes del mercado de "automóviles autónomos" son Tesla Motors, General Motors, Volvo y Mercedes-Benz. Corporaciones tales como Waymo, la hermana de Google, y Uber también están diseñando y probando la tecnología.

El camino hacia una autonomía completa en la que se use esta tecnología es complicado. Se han producido muchos choques de alto perfil y accidentes graves que implicaron automóviles automáticos. Algunos estados de los EE. UU. ya han aprobado el uso limitado de los automóviles con piloto automático pero los investigadores consideran que se deberán fabricar algunos automóviles más antes de que la tecnología se popularice.

Una vez que la tecnología esté comprobada, ¿qué sigue? ¿Camiones automáticos? ¿Aviones? ¿Trenes?

¿Cual es el dispositivo que detecta o mide un evento?

Un sensor detecta cambios en el entorno y convierte el evento físico en una señal analógica o digital medible para su visualización o procesamiento posterior.

La transformación digital fomenta la innovación empresarial.

Los datos analizados pueden ser útiles para empresas y gobiernos.

Si se hace adecuadamente, ¿los dispositivos inteligentes pueden aprender y modificar su propio código en función de nuevos parámetros?

Cuando los dispositivos inteligentes se programan correctamente, pueden evaluar los datos disponibles y modificar procesos o configuraciones en función de nuevos parámetros.

**1.2.1 La red es la base**

Los métodos que utilizamos para comunicarnos continúan evolucionando. Si bien una vez estuvimos limitados por cables y enchufes, las innovaciones en tecnología digital e inalámbrica han ampliado considerablemente el alcance de nuestras comunicaciones.

Las redes componen la base del mundo digitalizado. Hay redes de todo tamaño. Pueden ir desde redes simples, compuestas por dos PC, hasta redes que conectan millones de dispositivos.

Las redes domésticas simples permiten la conectividad a Internet. También permiten compartir recursos, por ejemplo: impresoras, documentos, imágenes y música, entre algunas computadoras locales.

En empresas y grandes organizaciones, las redes pueden proporcionar productos y servicios a los clientes a través de su conexión a Internet. Las redes también se pueden usar en una escala aún más amplia para proporcionar consolidación, almacenamiento y acceso a la información en los servidores de la red. Las redes permiten el correo electrónico, la mensajería instantánea y la colaboración entre los empleados. Además, la red permite la conectividad a nuevos lugares, lo que da más valor a las máquinas en entornos industriales.

Internet es la red más grande que existe, y proporciona “la capa electrónica” que rodea al planeta de manera eficaz. De hecho, el término Internet significa “red de redes”. Internet es, literalmente, una colección de redes privadas y públicas interconectadas. Las empresas, las redes de oficinas pequeñas y las redes domésticas se conectan a Internet.

1.2.2 Tipos de red

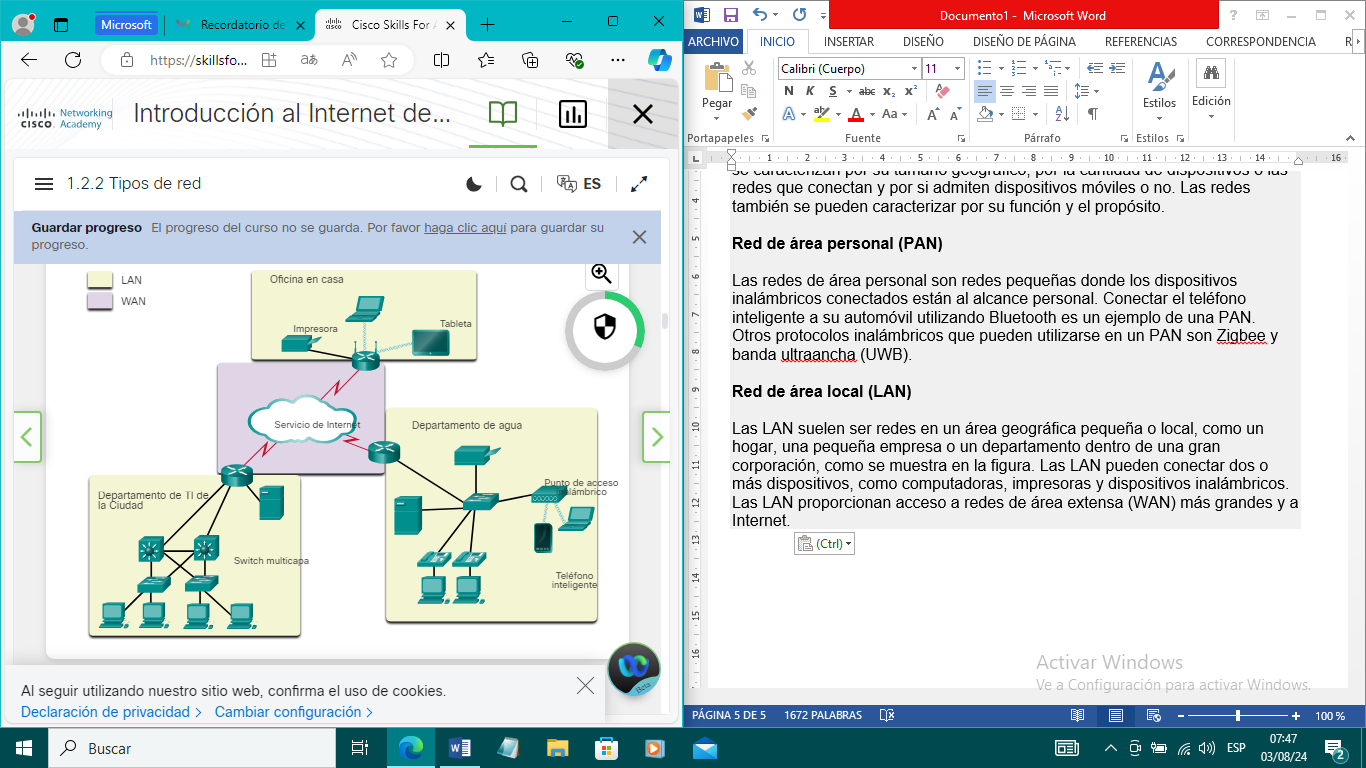
Las redes modernas pueden ser un poco confusas. Existen muchos tipos que se caracterizan por su tamaño geográfico, por la cantidad de dispositivos o las redes que conectan y por si admiten dispositivos móviles o no. Las redes también se pueden caracterizar por su función y el propósito.

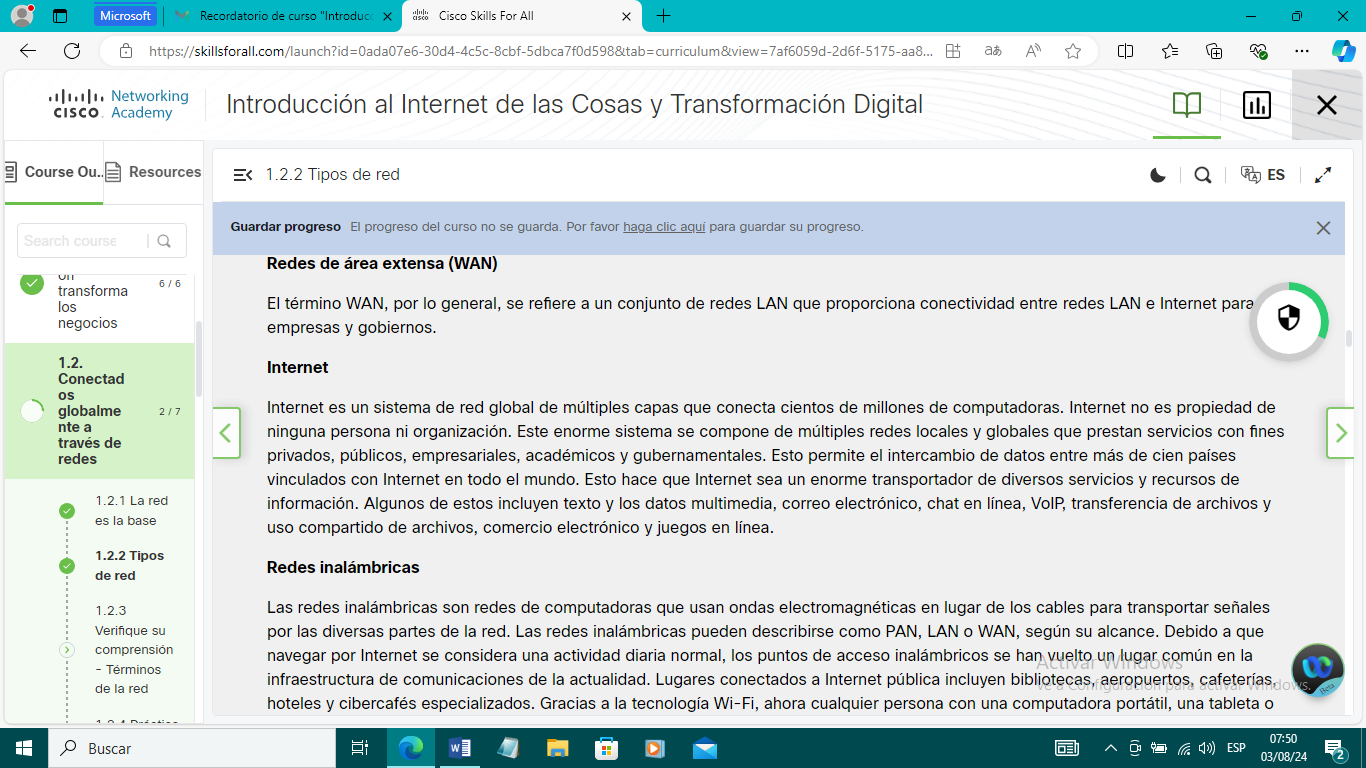
**Red de área personal (PAN)**

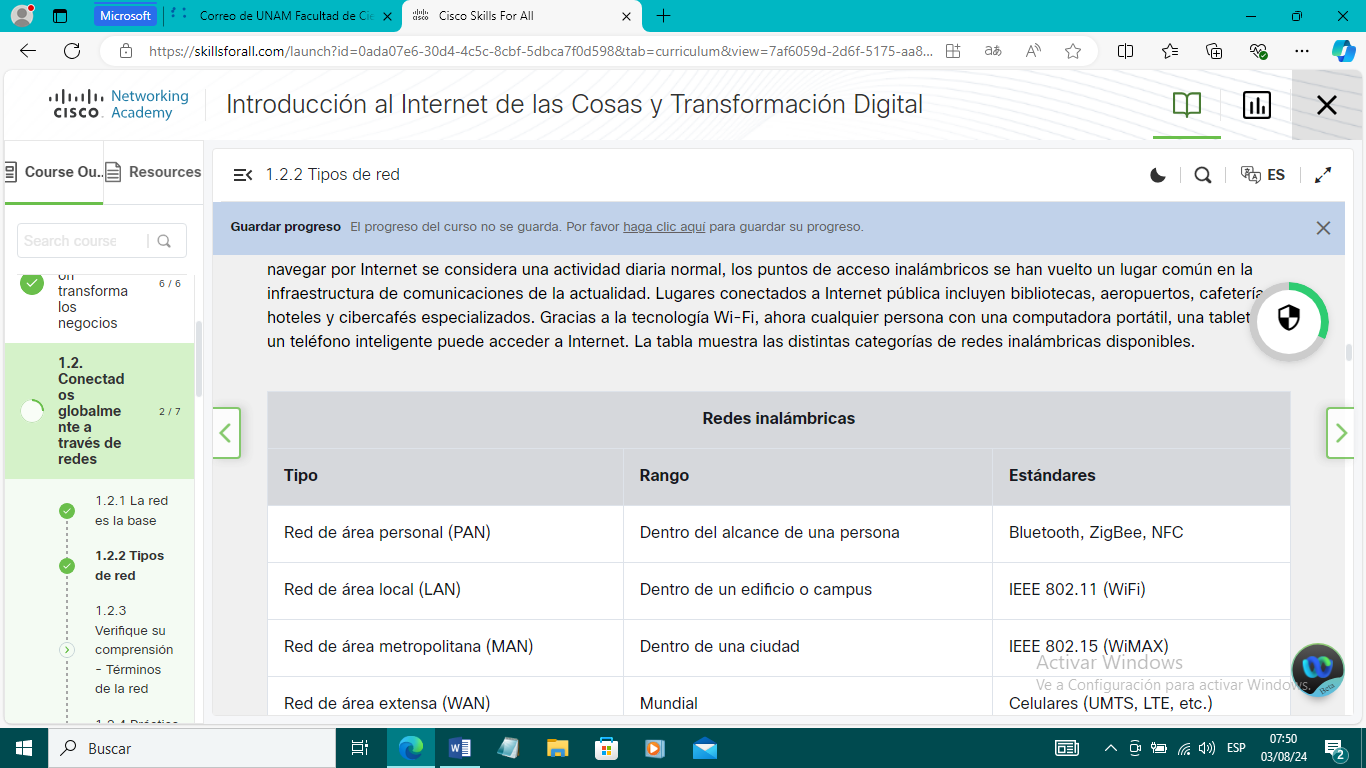
Las redes de área personal son redes pequeñas donde los dispositivos inalámbricos conectados están al alcance personal. Conectar el teléfono inteligente a su automóvil utilizando Bluetooth es un ejemplo de una PAN. Otros protocolos inalámbricos que pueden utilizarse en un PAN son Zigbee y banda ultraancha (UWB).

**Red de área local (LAN)**

Las LAN suelen ser redes en un área geográfica pequeña o local, como un hogar, una pequeña empresa o un departamento dentro de una gran corporación, como se muestra en la figura. Las LAN pueden conectar dos o más dispositivos, como computadoras, impresoras y dispositivos inalámbricos. Las LAN proporcionan acceso a redes de área extensa (WAN) más grandes y a Internet.







**La nube**

El término "en la nube" se usa en muchas formas diferentes. La nube no es tanto un tipo de red sino una colección de centros de datos o grupos de servidores conectados que se utilizan para almacenar y analizar datos, proporcionar acceso a aplicaciones en línea y proporcionar servicios de respaldo para uso personal y corporativo. Los servicios en la nube son proporcionados por diferentes organizaciones.

**Perímetro**

El borde o perímetro se refiere al límite entre la red de una organización e Internet. Define el lugar en el que una red administrada por una organización se conecta a una red administrada por otra organización. Un ejemplo común es el punto en el que una red corporativa se conecta a Internet. Una organización puede tener más de un perímetro de red, según su tamaño y la ubicación de sus instalaciones.

La computación en el borde permite que los datos locales se pre-procesen en el borde o perímetro de la red.

Bluetooth permite que los dispositivos inalámbricos se conecten a otros dispositivos a una distancia del brazo de una persona.

Las redes inalámbricas utilizan ondas electromagnéticas para transportar señales a través de las distintas partes de la red.

La nube es una colección de servidores y centros de datos para proporcionar almacenamiento de datos y acceso a datos desde cualquier lugar, en cualquier momento y en cualquier dispositivo.

LAN es un término que describe la red de una pequeña empresa que podría conectar 3 PC y 4 dispositivos inalámbricos.

1.3.1 ¿Qué es IoT?

Internet de las cosas (IoT) es la conexión de millones de dispositivos inteligentes y sensores conectados a Internet. Estos dispositivos y sensores conectados recopilan y comparten datos para que muchas organizaciones los usen y evalúen. Estas organizaciones incluyen las empresas, ciudades, gobiernos, hospitales y personas. IoT ha sido posible, en parte, debido a la llegada de procesadores asequibles y redes inalámbricas. Los objetos previamente inanimados, como picaportes o lámparas, ahora pueden contar con un sensor inteligente que puede recopilar y transferir datos a una red.

Los investigadores estiman que habrá 38.6 billones de dispositivos de IoT conectados a Internet en 2025 y 50 billones para 2030. ¡Son 190 millones de nuevos dispositivos conectados cada mes!

Quizás un tercio de los dispositivos conectados serán computadoras, teléfonos inteligentes, tabletas y televisores inteligentes. Los restantes dos tercios serán otros tipos de las "cosas": sensores, accionadores y dispositivos inteligentes recientemente inventados que supervisen, controlen, analicen y optimicen el mundo.

Algunos ejemplos de sensores conectados inteligentes son: timbres inteligentes, puertas de garaje, termostatos, equipos portátiles deportivos, marcapasos, semáforos, lugares de estacionamiento y muchos más. El límite de los diferentes objetos que podrían volverse sensores inteligentes está fijado solo por nuestra imaginación.

1.3.5 ¿Cómo se conectan los dispositivos IoT a la red?

Un sensor deben estar conectado a una red para que los datos recopilados puedan guardarse y compartirse. Esto requiere una conexión Ethernet cableada o una conexión inalámbrica a un controlador. Si bien se puede utilizar Ethernet inalámbrica, las alternativas de baja potencia como Bluetooth LE, Zigbee o LoRa son más viables. Los controladores son responsables de recopilar datos de los sensores y proporcionar conectividad hacia la red o Internet. Los controladores pueden tener la capacidad de tomar decisiones inmediatas o de enviar datos a una computadora más potente para su análisis. Esta computadora más potente puede estar en la misma LAN que el controlador, o bien puede ser accesible únicamente por medio de una conexión a Internet.

Los sensores a menudo funcionan junto con un dispositivo denominado actuador. Los actuadores toman la entrada eléctrica y transforman la entrada en acción física. Por ejemplo, si un sensor detecta calor excesivo en una sala, el sensor envía la lectura de temperatura al microcontrolador. El microcontrolador puede enviar los datos a un actuador, que a su vez encendería el aire acondicionado.

La mayoría de los nuevos dispositivos, tales como equipos portátiles para entrenar, marcapasos implantados, medidores de aire en una mina y medidores de agua en el campo de un establecimiento agrícola, requieren conectividad inalámbrica. Debido a que muchos sensores están "en el campo" y son impulsados por baterías o paneles solares, se debe tener en cuenta el consumo de energía. Se deben utilizar opciones de conexión de baja potencia para optim.9izar y ampliar la disponibilidad del sensor.

Las empresas que crean dispositivos portátiles poseen mucha información personal acerca de los usuarios ☹

Los gobiernos pueden monitorear el medioambiente. ☺

La dependencia de las compras en línea puede costar puestos de trabajo ☹

Los fabricantes pueden reducir el tiempo de inactividad mediante la predicción de los requisitos del mantenimiento. ☺

Los minoristas conocen todo lo que usted compra. ☹

Las ciudades pueden evaluar los futuros requisitos de transporte. ☺

Una falla de red puede ser catastrófica. ☹

Las personas pueden reducir los costos de energía residencial y los sistemas de calefacción ☺

.4.5 Video - Redes Basadas en la Intención de Cisco

En la actualidad, las redes conectan a miles de millones de sensores. A través del software, los datos de estos sensores pueden realizar cambios en entornos físicos sin intervención humana.

Como se mencionó anteriormente, todos los dispositivos digitales funcionan según los programas informáticos y los datos suministrados. La inteligencia artificial implica que estos dispositivos son capaces de pensar por sí mismos. Si se programan de manera correcta, los dispositivos inteligentes pueden evaluar los datos que reciben, y modificar los procesos o la configuración de inmediato. Si se les proporcionan los datos suficientes, pueden “aprender” y modificar sus propios códigos según los nuevos parámetros.

¿Qué es lo siguiente?

Sabemos que se puede escribir software para permitir que los datos modifiquen parámetros dentro del código a fin de cambiar la configuración de la temperatura de su hogar o la velocidad con la que su hijo adolescente puede conducir el automóvil familiar. ¿Por qué no ofreceríamos software con reglas, pautas o intenciones, de modo que los datos pudieran modificar la red, las funciones de la infraestructura o las funciones de seguridad dentro de una red? Esto es, en realidad, ya posible. Se denomina Redes basadas en la intención (IBN).

El siguiente es un ejemplo simple para comprender mejor el concepto de IBN: es posible que la empresa establezca que un empleado contratado reciba acceso a solo un conjunto específico de datos y aplicaciones. Esta es la **intención**. En un sistema de redes basadas en la intención (IBN), todos los dispositivos de red se configurarán automáticamente para cumplir con este requisito en toda la red, sin importar dónde esté conectado el empleado. La VLAN, la subred, la ACL y todos los otros detalles se definen y configuran automáticamente siguiendo las mejores prácticas. La intención se debe definir una sola vez en una ⁪consola de administración central y luego la red la ratificará continuamente, incluso si hay cambios en la red.

1.5.1 Resumen de Todo está Conectado

expand\_less

La digitalización transforma los negocios

El mundo rápidamente se cubre por redes que permiten que los dispositivos digitales se interconecten y transmitan. Conforme las redes digitales sigan creciendo en todo el mundo y conforme los beneficios económicos de la digitalización sigan aumentando, seremos testigos de una transformación digital. La transformación digital es la aplicación de la tecnología digital para proporcionar el entorno adecuado para la innovación de las empresas y la industria.

Los sensores ahora están en todas partes, y recopilan y transmiten cantidades enormes de datos. Los datos generados pueden almacenarse y analizarse en una fecha posterior, o se pueden analizar y utilizar inmediatamente. Los datos analizados son usados por gobiernos, ciudades, empresas y personas para generar cambios, como el monitoreo del medioambiente, el pronóstico del crecimiento de la población, el control de residuos o la protección de una vivienda.

Conectados globalmente a través de redes

Las redes componen la base del mundo digitalizado. Existen muchos tipos de redes que se caracterizan por su tamaño geográfico, por la cantidad de dispositivos o las redes que se conectan y por si admiten dispositivos móviles o no. Las redes también se pueden caracterizar por su función y el propósito.

* PAN: Bluetooth
* LAN
* WAN: Internet, la nube, computación en el borde
* Tecnología inalámbrica: Wi-Fi, red celular

El crecimiento de los dispositivos de IoT

Por lo general, un sensor se conecta a un controlador mediante una conexión inalámbrica. Los controladores recopilan los datos de los sensores, y envían los datos para su almacenamiento o análisis. Los controladores pueden tener la capacidad para tomar decisiones inmediatas o pueden funcionar junto con un dispositivo denominado accionador. Los accionadores toman la entrada eléctrica y transforman la entrada en acción física.

Conexión de los dispositivos de IoT a la red

En la actualidad, las redes conectan a miles de millones de sensores y tienen la capacidad para realizar cambios en entornos físicos sin intervención humana. El futuro de las redes girará en torno a la inteligencia artificial (AI) y las redes basadas en la intención (IBN). Si se programan de manera correcta, los dispositivos inteligentes pueden evaluar los datos que reciben, y modificar los procesos o la configuración. Si se les proporcionan los datos suficientes, pueden “aprender” y modificar sus propios códigos según los nuevos parámetros.

2.1.2 Diagramas de flujo

Los diagramas de flujo se utilizan en diferentes industrias, entre ellas ingeniería, ciencias físicas y programación de computadoras, en los cuales se requiere una comprensión completa de los procesos o flujos de trabajo. Los diagramas de flujo son diagramas que se utilizan para representar estos procesos o flujos de trabajo.

Los diagramas de flujo ilustran cómo debe funcionar un proceso. Los diagramas de flujo no requieren terminología ni símbolos complejos y específicos de la industria. Un diagrama de flujo debe ser fácil de entender sin tener que volverse un experto en el campo elegido.

Los diagramas de flujo deben mostrar los estados de entrada, las decisiones tomadas y los resultados de dichas decisiones. Es importante mostrar los pasos que se deben seguir cuando el resultado de una decisión es sí o no.

Es común que los programadores creen el primer borrador de un programa sin ningún lenguaje de programación específico. Estos programas independientes del lenguaje se centran en la lógica en lugar de la sintaxis y se conocen como algoritmos. Un diagrama de flujo es una manera común de representar un algoritmo. En la figura, se muestra un ejemplo de diagrama de flujo.

2.1.3 Software del sistema, software de aplicaciones y lenguajes informáticos

Existen dos tipos comunes de software informático: software del sistema y software de aplicaciones.

Los programas del software de aplicaciones se crean con el fin de realizar una tarea determinada o un conjunto de tareas. Por ejemplo, Cisco Packet Tracer es un programa de simulación de red que permite que los usuarios modelen redes complejas y formulen preguntas hipotéticas sobre el comportamiento de la red.

El software del sistema funciona entre el hardware de la computadora y el programa de aplicaciones. Es el software del sistema que controla el hardware de la computadora y permite que los programas de las aplicaciones funcionen. Entre los ejemplos comunes de software del sistema se incluyen Linux, Apple OSX, y Microsoft Windows.

El software del sistema y el software de aplicaciones se crean con un lenguaje de programación. Un lenguaje de programación es un lenguaje formal diseñado para crear programas que comunican instrucciones al hardware de la computadora. Estos programas implementan algoritmos que son conjuntos autónomos y pormenorizados de las operaciones que se deben realizar.

Algunos lenguajes informáticos compilan sus programas en un conjunto de instrucciones de lenguaje de máquina. C++ es un ejemplo de lenguaje informático compilado. Otros interpretan estas instrucciones directamente sin primero compilarlos en un lenguaje de máquina. Python es un ejemplo de un lenguaje de programación interpretado. En la figura, se muestra un ejemplo de código de Python.

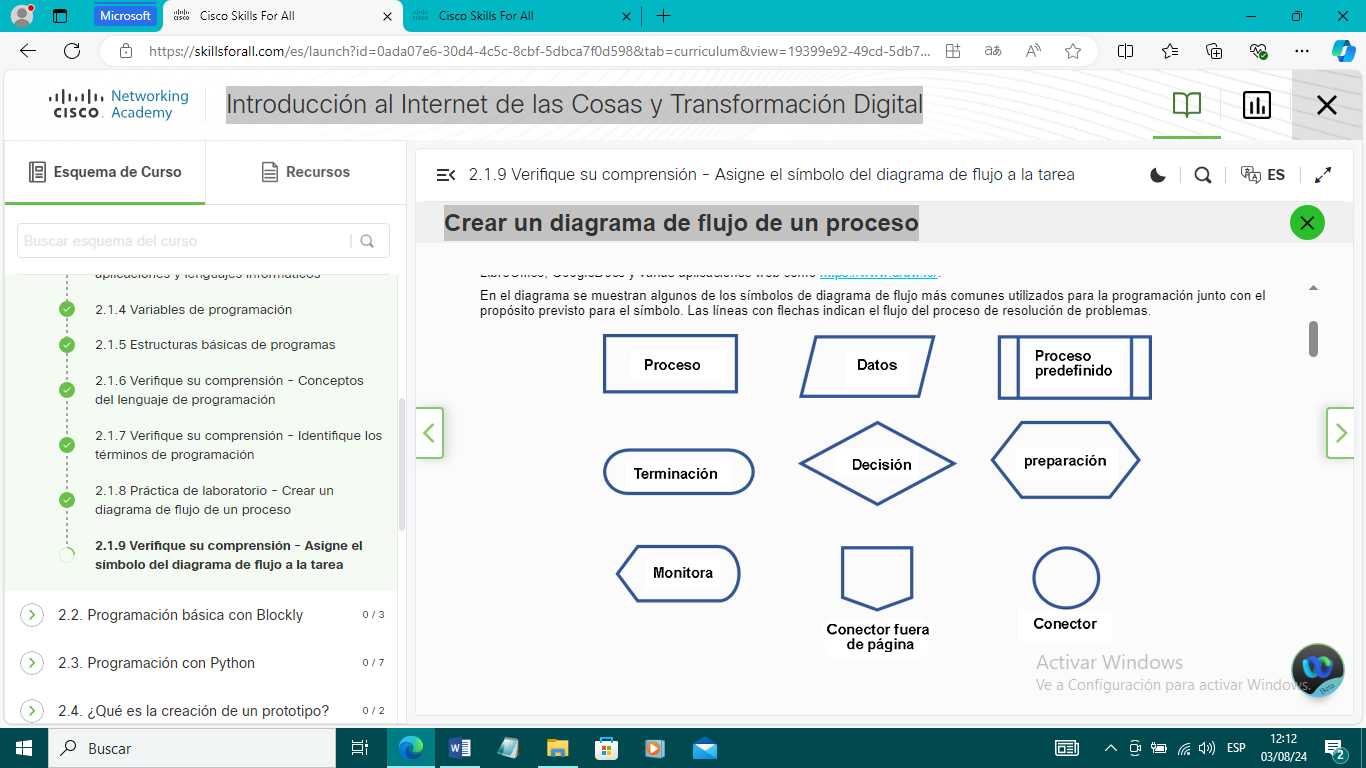
Cuando se determina el lenguaje de programación y se diagrama el proceso en un diagrama de flujo, puede comenzar la creación del programa. La mayoría de los lenguajes informáticos usan estructuras de programa similares.

**Crear un diagrama de flujo de un proceso**

Los diagramas de flujo son diagramas utilizados para representar procesos o flujos de trabajo. Haciendo uso de diferentes formas, cuadros y flechas de conexión, el diagrama de flujo representa el flujo de la solución a un problema determinado. Los diagramas de flujo se usan comúnmente para representar programas, algoritmos o cualquier proceso pedido en varias disciplinas. Normalmente, los diagramas de flujo se elaboran antes de comenzar un proceso o desarrollar una aplicación con el fin de verificar y detectar posibles flujos lógicos hacia la solución antes de que se desarrolle e implemente.

Los diagramas de flujo se pueden dibujar o crear a mano utilizando una serie de paquetes, incluidos los productos de Microsoft Office, LibreOffice, GoogleDocs y varias aplicaciones web como <https://www.draw.io/>.

En el diagrama se muestran algunos de los símbolos de diagrama de flujo más comunes utilizados para la programación junto con el propósito previsto para el símbolo. Las líneas con flechas indican el flujo del proceso de resolución de problemas.



Google proporciona una serie de juegos educativos gratuitos y de código abierto que pueden ayudarlo a aprender a programar. La serie se denomina Juegos de Blockly.

Para obtener más información sobre Blockly Games o para probarlo usted mismo, vaya a [https://blockly.games](https://blockly.games/)

3.3.1 ¿Por qué las empresas analizan datos?

Cada organización debe volverse más eficiente y más innovadoras para mantenerse competitiva y relevante en el mundo digitalizado. IoT es una parte integral para lograr esa eficiencia e innovación.

El objetivo de muchas empresas es recopilar y analizar cantidades masivas de nuevos datos sobre el uso de productos y obtener conocimientos valiosos. El análisis de datos permite que las empresas comprendan mejor el impacto de sus productos y servicios, ajusten sus métodos y objetivos, y proporcionen a sus clientes mejores productos más rápido. La capacidad para obtener nuevas perspectivas a partir de los datos aporta valor a la empresa.

Para los negocios, los datos son el nuevo petróleo. Como el petróleo crudo, tiene valor, pero si está sin refinar no puede utilizarse con facilidad. El petróleo crudo debe transformarse en gasolina, plástico, sustancias químicas y otras sustancias para que sea un producto valioso. Es lo mismo con los datos. Los datos se deben descomponer y analizar para que tengan valor.

Los valores provienen de los dos tipos de datos procesados principales: transaccionales y analíticos. La información transaccional se captura y se procesa a medida que se producen eventos. La información transaccional se utiliza para analizar informes de ventas y planes de fabricación diarios a fin de determinar cuánto inventario transportar. La información analítica permite que se realicen tareas de análisis a nivel gerencial, como determinar si la organización debe instalar una nueva planta de fabricación o contratar personal de ventas adicional.

3.3.2 Fuentes de información

La fuente de los datos en los grandes conjuntos de datos es variada. Además de los datos de los sensores, otros datos se originan a partir de todo lo que se ha analizado, introducido y divulgado en Internet de fuentes tales como las siguientes:

* Redes sociales - Facebook, YouTube, WhatsApp, WeChat, TikTok e Instagram
* HTTP, páginas web y motores de búsqueda en Internet
* Datos históricos de archivos públicos y privados
* Metadatos que se adjuntan a correos electrónicos, imágenes y documentos transmitidos
* Formularios médicos, formularios de seguros y formularios de impuestos
* Investigación genómica mediante ADN

Los datos recopilados se pueden clasificar como estructurados o no estructurados.

Los datos estructurados se crean mediante aplicaciones que utilizan entradas de formato “fijo”, como hojas de cálculo o formularios médicos. Incluso si los datos se consideran estructurados, diferentes aplicaciones crean archivos en distintos formatos que no necesariamente son compatibles unos con otros. Es posible que los datos estructurados se deban manipular en un formato común como CSV.

Los archivos de valores separados por comas (CSV) son un tipo de archivo de texto simple que usa comas para separar columnas en una tabla de datos, y el carácter de retorno de párrafo para separar filas. Cada fila es un registro. A pesar de ser comúnmente utilizados para importar y exportar en bases de datos y hojas de cálculo tradicionales, no existe un estándar específico. JSON y XML también son tipos de archivo de texto simple que utilizan un modo estándar para representar registros de datos. Estos formatos de archivo son compatibles con una amplia gama de aplicaciones. La conversión de datos en un formato común es una manera valiosa de combinar datos de diferentes orígenes.

Los datos no estructurados se generan en un estilo de “forma libre”, como audio, video, páginas web y tweets. Los datos no estructurados requieren diferentes herramientas para preparar los datos para análisis o procesamiento. Los siguientes son dos ejemplos:

* Los sitios web se crean para proporcionar datos a los seres humanos, no a las máquinas. Las herramientas que "raspan la red" (web scraping) extraen datos de páginas HTML automáticamente. Esto es similar al rastreador o araña web de un motor de búsqueda. Explora la web para extraer datos y crear bases de datos para responder a las consultas de búsqueda. El software de web scraping puede utilizar el Protocolo de transferencia de hipertexto o un explorador web para acceder a la World Wide Web. Por lo general, el proceso de web scraping es un proceso automatizado que utiliza un bot o un rastreador web para obtener datos. Los datos específicos se recopilan y se copian de la web a una base de datos o una hoja de cálculo. Los datos pueden luego analizarse fácilmente.
* Muchos proveedores de servicios web importantes como Facebook proporcionan interfaces estandarizadas para recopilar los datos automáticamente mediante interfaces de programación de aplicaciones (API). El enfoque más habitual es usar API RESTful. Las API RESTful usan HTTP como el protocolo de comunicación y estructura de JSON para codificar los datos. Las páginas web de Internet como Google y Twitter recopilan grandes cantidades de datos estáticos y de series de tiempo. El conocimiento de API para estos sitios permite que los analistas y los ingenieros de datos tengan acceso a las grandes cantidades de datos que se generan constantemente en Internet.

3.3.3 Visualización de datos

La minería de datos es el proceso que se utiliza para convertir los datos sin procesar en información significativa al detectar patrones y relaciones en los grandes conjuntos de datos.

Para ser de valor, los datos sometidos a minería de datos se deben analizar y presentar a los administradores y las personas responsables de tomar decisiones. Existen muchas visualizaciones de diferentes que se pueden utilizar para presentar el valor de los datos. La determinación del mejor gráfico a utilizar variará según lo siguiente:

* Cantidad de variables para mostrar
* Cantidad de puntos de datos en cada variable
* Representan los datos una línea de tiempo
* Los elementos requieren comparaciones

Algunos de los tipos más comunes de gráficos son circulares, de líneas, de columnas, de barras y de dispersión.

3.3.6 Análisis de datos masivos para el uso eficaz en la empresa

Los datos masivos son solo eso: ¡GRANDES! Resultan más útiles si puede analizarlos para obtener valor a partir de ellos. El análisis de datos es el proceso de inspección, limpieza, transformación y creación de modelos de datos para descubrir información útil. El análisis de los datos masivos generalmente requiere herramientas y aplicaciones creadas para este propósito. Estas herramientas de análisis se diseñaron para brindar a las empresas conocimientos valiosos, patrones e información detallada.

Antes de comenzar cualquier análisis, es fundamental saber qué problema de la empresa intenta resolver o qué información de la empresa busca. ¿Se interesan por el comportamiento del cliente en estados específicos, los patrones de consumo de energía en cuadrantes de la ciudad diferentes o la cantidad de "me gusta" de Facebook en función de la edad?

Tener una estrategia permite que una empresa determine el tipo de análisis requerido y la mejor herramienta para realizar el análisis. Una estrategia también ayuda a determinar la manera más eficaz de presentar los resultados para la administración.

Las herramientas y aplicaciones varían desde el uso de una hoja de cálculo de Excel o Google Analytics para muestras de datos de pequeñas a medianas, hasta las aplicaciones dedicadas a la manipulación y al análisis de conjuntos de datos realmente masivos.

Hay muchas herramientas de escritorio para el análisis de datos masivos que las empresas pueden seleccionar, como: Knime, OpenRefine, Orange y RapidMiner. Las herramientas de análisis basadas en la nube incluyen Big Query de Google, IBM Cognos Analytics, TIBCO Spotfire y Board.