

## ***Big Data y el camino a la inteligencia artificial autonómica***

### ***Big Data and the road towards and autonomic artificial intelligence***

**Victor Manuel Castillo Araujo**

Universidad Nacional Experimental Politécnica "Antonio José de Sucre". Caracas, Venezuela.

Teléfono: +58-414-1375149, victorc25@gmail.com,

Ponencia adscrita a la Línea de Investigación Doctoral "Tecnología e Innovación"

Sub-Tema del evento: Combatir el VIH/SIDA, el paludismo y otras enfermedades.

#### **RESUMEN**

Al día de hoy, diferentes disciplinas y tecnologías han sido desarrolladas para responder a una gran variedad de necesidades de la sociedad. Una de las tendencias más importantes de los últimos tiempos se refiere a la convergencia de la tecnología y la capacidad de poder procesar una cantidad de datos que previamente se hubiese considerado impensable. Una de estas tecnologías es llamada *Big Data* y a través de este artículo y una revisión teórica de otras que la complementan, se presenta una alternativa de cómo pueden evolucionar sus conceptos y crear un sistema que es capaz de analizar de manera autonómica patrones y tendencias y tomar o proponer acciones para resolver diferentes situaciones, entre ellas, enfermedades y distribución de los recursos en la sociedad.

**Palabras claves:** Grandes Datos, Inteligencia Artificial, Aprendizaje de Máquina, Plataforma Definida por Software, Internet de las Cosas

**Keywords:** Big Data, Artificial Intelligence, Machine Learning, Software Defined Platform, Internet of Things

# ***Big Data y el camino a la inteligencia artificial autónoma***

## ***Big Data and the road towards an autonomic artificial intelligence***

**Victor Manuel Castillo Araujo**

Universidad Nacional Experimental Politécnica "Antonio José de Sucre". Caracas, Venezuela.

Teléfono: +58-414-1375149, victorc25@gmail.com,

Ponencia adscrita a la Línea de Investigación Doctoral "Tecnología e Innovación"

Sub-Tema del evento: Combatir el VIH/SIDA, el paludismo y otras enfermedades.

### **1.- Introducción**

Día a día se genera mucha más información de la que se puede procesar con los mecanismos tradicionales, por lo que mucha información muy valiosa y que puede aportar de manera positiva a la sociedad se pierde por las limitaciones de esos mecanismos. La tecnología y la ciencia han avanzado mucho en muchas disciplinas diferentes y, siguiendo una de las tendencias con más fuerza en la actualidad en el mundo de la tecnología, nos encontramos en un punto de inflexión donde todas esas disciplinas terminarán convergiendo, permitiendo no sólo el procesamiento de toda la información, sino también, avanzar hacia una inteligencia aumentada que facilitará a la sociedad tomar decisiones para el mejoramiento de todas las condiciones humanas.

Esta tendencia es a su vez un gran reto. Lograr integrar, no sólo la tecnología, sino también el conocimiento y las habilidades hacia un objetivo común en equipos multidisciplinarios implica un cambio en la cultura y la manera de trabajar con un solapamiento de todas las ciencias.

Este artículo presenta, con un foco tecnológico, una manera de cómo puede evolucionar la cultura y la tecnología en si para atender diversos temas, entre los cuales se encuentran tanto las enfermedades, sean infecciosas, genéticas o inducidas por el medio, como la distribución de los recursos y riqueza en la sociedad para permitir el crecimiento de todos sus individuos.

Parte de la tecnología planteada en esta convergencia va desde algunas que se encuentran en su punto más alto de exposición, como *Big Data* e Internet de las Cosas, hasta otras que han venido siendo desarrolladas y mejoradas a lo largo de un tiempo muy extenso, como lo es la Inteligencia Artificial y el Aprendizaje de Máquina.

Se planteará a través de un desarrollo lógico cómo estas piezas se pueden ver como partes de un rompecabezas y vienen a integrarse en algo que va mucho más allá de las capacidades que pueden ofrecer cada una por separado y como esta convergencia permitirá resolver problemas que, en algunos casos, hoy en día consideramos muy difíciles, por no decir imposibles de resolver.

## 2.- Objetivo

El presente artículo busca hacer una revisión de cómo la tecnología actual y las tendencias en el medio pueden ser utilizadas para resolver dificultades que existen en la sociedad, por medio de la integración de la misma y el foco en resolver algunos de los problemas establecidos dentro de las Metas del Milenio.

## 3.- Materiales y Métodos

A fin de realizar este artículo se realizó una investigación en la literatura científica y de industria sobre las diferentes disciplinas y tendencias tecnológicas y su aplicabilidad en distintos ámbitos de la vida humana. Para la investigación, después de recopilar la información requerida, se procedió a realizar el análisis y la comprensión de las necesidades que enfrentamos como sociedad para mejorar la salud y las condiciones de vida. Se presenta al lector con una descripción cualitativa de como la tecnología se podría convertir en un habilitador para esta meta, recorriendo los conceptos base y el análisis de la aplicabilidad de los mismos en la sociedad.

## 4.- Conceptos

El término **inteligencia** proviene del latín *intelligentia* (entendimiento, capacidad para entender), que a su vez deriva de *intelligere*. Esta es una palabra compuesta por otros dos términos: inter (“entre”) y legere (“escoger, seleccionar”). Por lo tanto, el origen etimológico del concepto de inteligencia hace referencia a quien sabe elegir. La inteligencia posibilita la selección de las alternativas más convenientes para la resolución de un problema, basado en conocimiento o experiencias previas. De acuerdo a lo descrito en la etimología, un individuo es inteligente cuando es capaz de escoger la mejor opción entre las posibilidades que se presentan a su alcance para resolver un problema.

La inteligencia está, por lo tanto, relacionada a la habilidad de un individuo de captar información, comprenderla, procesarla y emplearla de manera acertada. Quiere decir que es la capacidad de relacionar conocimientos y conceptos que permiten la resolución de un determinado conflicto, es una cualidad que poseemos naturalmente los humanos y también los animales, sólo que en el caso de los últimos se le llama simplemente instinto.

La inteligencia es la capacidad de asimilar, guardar, elaborar información y utilizarla para resolver problemas, cosa que también son capaces de hacer los animales e incluso los ordenadores. Pero el ser humano va más allá, desarrollando una capacidad de iniciar, dirigir y controlar nuestras operaciones mentales y todas las actividades que manejan información. Aprendemos, reconocemos, relacionamos, mantenemos el

equilibrio y muchas cosas más sin saber cómo lo hacemos, lo cual se identifica como el sistema nervioso autónomo o vegetativo. Pero tenemos además pensamiento abstracto y lógica que nos brindan la capacidad de integrar estas actividades mentales y de hacerlas voluntarias, en definitiva de controlarlas, como ocurre con nuestra atención o con el aprendizaje, que deja de ser automático como en los animales para focalizarlo hacia determinados objetivos deseados.

La función principal de la inteligencia no es sólo conocer, sino dirigir el comportamiento para resolver problemas de la vida cotidiana con eficacia.

El concepto de **Inteligencia Artificial** fue desarrollado para referirse a los sistemas creados por los seres humanos que son capaces de planificar, elaborar pensamientos abstractos, entender ideas y aprender. La Inteligencia Artificial se origina en la fusión de la fisiología y la filosofía y consiste en darle cierto tipo de vida inteligente a un objeto carente naturalmente de ella, que vino a ser complementada en la edad moderna con las ciencias computacionales. Se desarrolla a partir de códigos de lenguaje de programación y de acuerdo a la complejidad de los mismos se pueden crear entes más o menos inteligentes, capaces de tomar buenas, mediocres o malas decisiones. Es la inteligencia que poseen hoy en día los robots, máquinas, software y los aparatos con cierta independencia.

Esta inteligencia les proporciona a las máquinas la capacidad de percibir variables en su medio y tomar acciones con algún fin. Esas acciones y el resultado dependen tradicionalmente de la complejidad con la que fue diseñada la inteligencia, la capacidad de procesar eventos e información y la extensión de las acciones que puede tomar en sí.

*“La inteligencia consiste no sólo en el conocimiento, sino también en la destreza de aplicar los conocimientos en la práctica.”*

Aristóteles (384 AC-322 AC) Filósofo griego.

Tradicionalmente, la información con la que cuenta una Inteligencia Artificial es limitada a las variables para la que fue diseñada y puede procesar y, en general, son limitadas a realizar una única función. La incertidumbre es manejada en el reino de las estadísticas y probabilidades.

El **Aprendizaje de Máquina** (Machine Learning) es una disciplina científica que nació a partir de investigaciones en Inteligencia Artificial y explora el estudio de algoritmos que pueden aprender de los datos, construyendo modelos en base a ejemplos y utilizándolos para tomar decisiones o hacer predicciones, siguiendo instrucciones programadas. El Aprendizaje de Máquina es afín a los estudios en estadísticas y optimizaciones matemáticas.

El Aprendizaje de Máquina también le proporciona a un sistema la capacidad de predicción, por medio de los modelos que desarrolla al generalizar de sus experiencias, por lo que puede tomar decisiones preventivas para modificar los resultados, incluso durante el desarrollo de los eventos.

Un factor muy importante que afecta los resultados a los que una inteligencia puede llegar es el de información que utiliza y como la procesa. Una restricción de la Inteligencia Artificial ha sido entorno a la cantidad, velocidad y tipos de datos que puede utilizar como insumo para tomar decisiones. Una inteligencia natural es capaz de procesar información proveniente de diferentes sentidos, combinarlos con conocimientos y experiencias previas y llegar a conclusiones, por lo que una Inteligencia Artificial avanzada debe poseer un alcance similar. Recientemente ha tenido una explosión muy fuerte el campo de los grandes datos (**Big Data**), cuyos retos residen en la captura, análisis, almacenamiento, transferencia y visualización de datos de alto volumen, gran variedad y que se producen a una velocidad muy grande desde una diversidad muy grande de fuentes.

Una infraestructura de Big Data involucra una plataforma tecnológica de múltiples nodos de cómputo y almacenamiento distribuido, que puede realizar una gran cantidad de tareas en paralelo y organizar toda esta información para que pueda ser utilizada. Los datos son organizados, de modo que los Científicos de Datos (*Data Scientists*) puedan analizarlos posteriormente.

El concepto de “**Todo Definido por Software**” (Software Defined Everything o SDx) agrupa varios términos que en la industria tecnológica vienen a flexibilizar y facilitar la manera en que se administran y configuran los diferentes componentes de una plataforma (redes, almacenamiento, cómputo) y se convierte en un habilitador para la tendencia de **Cómputo en Nube**. Nacido del requerimiento de sistemas dinámicos y altamente variables en el tiempo, la capacidad de cambiar la configuración de una plataforma completa a través de software, sin requerir cambios físicos, abre las puertas a una tecnología completamente reprogramable. Una reconfiguración se traduce entonces en comandos y son cambios exclusivamente lógicos.

Con esta característica, un sistema moderno es capaz de modificar de una manera sencilla parámetros en su funcionamiento, para ajustarse a los requerimientos. Un ejemplo podría ser que cuando la cantidad de información a procesar aumenta, el sistema puede aumentar el cómputo y el ancho de banda de las redes, procesar el pico de información y una vez culminado, regresar a la configuración inicial. De este modo, los tiempos de respuesta del sistema pueden estar siempre muy cerca del tiempo real.

Un par de conceptos faltantes dentro del alcance de este artículo es el de **Internet de las Cosas** y **Tecnología Usable**. El primero se refiere a todos aquellos objetos físicos

relativamente comunes que han sido dotados de electrónica, software, sensores y conectividad para ofrecer más servicios y capacidades, comunicándose con otros dispositivos conectables. Estos dispositivos pueden ir desde sensores permanentemente conectados de diversos tipos en el hogar, industrias, automóviles o infraestructura de un país, entre otros, como accionadores que permiten cambiar variables en el medio y automatizar sistemas.

También se le otorga la definición de dispositivos inteligentes (Smart Devices) y la reducción física del volumen que ocupa la tecnología ha permitido que dichos dispositivos puedan ser llevados permanentemente por cada persona, empezando con los teléfonos inteligentes, ahora abriendo las puertas a relojes inteligentes, pero efectivamente haciéndose cada vez más comunes. La última tendencia al respecto es la Tecnología Usable enfocada a la salud y el bienestar de las personas.

Es tecnología que genera una cantidad muy grande de datos, a una velocidad muy alta, por lo que está generalmente relacionada con un procesamiento basado en Big Data.

## **5.- Análisis y Desarrollo**

Volviendo al concepto inicial de Inteligencia presentado en la sección anterior, se entiende que un punto clave para que las decisiones tomadas sean lo más óptima posible, se requiere de la mayor cantidad de información posible esté a disposición y pueda ser procesada. Una limitación de los sistemas de información actuales es que los datos producidos son de un volumen, velocidad y variedad tan alta que no es posible procesarlos, por lo que toda Inteligencia Artificial que se crea es, por concepto, limitada a esa capacidad de procesar los datos. Sin embargo, Big Data viene a modificar esta realidad, poniendo a disposición la información procesada para poder tomar decisiones.

Parece natural pensar que para permitir que la Inteligencia Artificial sea más similar a la inteligencia humana, es necesario proveer al sistema con más información y más capacidad para procesar esa información, a la vez que sea capaz de aprender de experiencias pasadas, pudiéndose pensar en una especie de retroalimentación de resultados previos como nueva información en base a la cual evaluar variables en situaciones similares. Es en este punto donde las tecnologías de grandes datos vienen a complementar el factor clave para potenciar la Inteligencia Artificial y viceversa.

La integración de Inteligencia Artificial y Big Data va incluso un paso más allá, partiendo del conocimiento de que, hasta ahora, para convertir los datos almacenados en una plataforma de Big Data en información valiosa, se requiere de especialistas, Científicos

de Datos que puedan extraer esta información, para poder procesarla y convertirla en accionables para tomar decisiones, con ciertos grados de automatización. La integración de Big Data con Inteligencia Artificial permitirá en un futuro próximo poder almacenar, procesar, tomar decisiones y actuar a las máquinas y software, para las funciones que estén preparadas. Sumado a esto, la capacidad de Aprendizaje de Máquina permitirá que el sistema pueda, en cierto modo, modificarse a sí mismo para tomar decisiones más acertadas, de acuerdo a la experiencia de decisiones previas, retroalimentando las potenciales acciones con esa información y considerando que los resultados en sí mismos son parte de la información que puede procesar.

Al considerar los resultados previos como datos para las decisiones futuras, es posible aumentar el conocimiento del sistema a no sólo datos fijos, sino adaptado al entorno y a la respuesta que el mismo da a las acciones. Esta condición implicaría que dos sistemas inicialmente idénticos de este tipo colocados en ambientes diferentes podrían llegar finalmente a tomar decisiones distintas, hecha la misma pregunta, asemejando a lo que sucede cuando dos personas se encuentran ante una misma situación, pero toman decisiones que pueden llegar a ser opuestas. Esto nos lleva a considerar la idea de sistemas especializados, no sólo por diseño, sino también por experiencias.

La expansión de la tendencia de todo definido por software a mucho más que sólo la plataforma tecnológica y su integración con los sistemas industriales y de distribución, puede permitir que parámetros de los mismos sean modificados sin requerir intervención humana para optimizar los resultados finales. Un ejemplo de esto son las guías de distribución de alimentos, que dependiendo del requerimiento de una región, el tráfico y el clima, pudiese estar cambiando su orden de despacho y rutas para garantizar la mayor eficiencia.

*“Las inteligencias poco capaces se interesan en lo extraordinario; las inteligencias poderosas en las cosas ordinarias.”*

Elbert Hubbard (1856-1915) Ensayista estadounidense.

Toda esta tecnología, avances y convergencia pueden aprovecharse especialmente en aquellos campos y fuentes de información cotidianos que tradicionalmente son descartados por el volumen y la complejidad de los datos, pero con un sistema que puede procesarlos, detectar patrones, analizarlos y tomar decisiones para actuar o dar sugerencias de acción puede transformar completamente el día a día para todos. Algunos ejemplos de la aplicación de esta tecnología planteada en temas cotidianos son:

- Detectar brotes de enfermedades contagiosas o no y definir las acciones necesarias para resolver estos brotes en las zonas afectadas. Puede ir desde

infecciones como el dengue, chikungunya, paludismo u otros hasta entender efectos de malnutrición en una región.

- Comunicación con humanos: detectar inflexiones y tonos de voz, entender idiomas y análisis de sentimiento, comprender comunicación verbal y no verbal.
- Entendimiento de los factores que afectan una sociedad. Capacidad para entender las noticias a través de los diferentes medios (Escritos, redes sociales, multimedios, entre otros) y prevenir comportamientos en la economía y la sociedad.
- Procesar datos ambientales históricos y presentes para mejorar los procesos relacionados a la agricultura y ganadería, previendo los efectos de las estaciones y temporales.
- Analizar variables del cuerpo humano para determinar si los patrones y parámetros se encuentran en los rangos correctos. Correlacionar con un historial médico familiar.
- Entender comportamientos en el tráfico terrestre y modificar automáticamente los sistemas como semáforos y rutas de transporte para optimizar el flujo.
- Modificar las condiciones de humedad, temperatura, posición de cortinas, flujo de aire y otras en el hogar, para ajustarse de la mejor manera a las personas que se encuentran en ella.
- Ejecutar tareas mecánicas sin intervención humana. Plantas e industrias que pueden operar con total autonomía, conociendo modificaciones estacionales de los procesos y aplicando los cambios correspondientes.
- Mantener un análisis permanente en los esquemas de distribución de alimentos y otros recursos y detectar comportamientos anormales.

El foco de este trabajo son los potenciales beneficios de aplicar la tecnología a los puntos relacionados a la atención de los temas de enfermedades, pobreza extrema y hambre y en los que se ahondará ahora con un poco más de profundidad.

El secreto para resolver los temas del bienestar de la sociedad son los datos. La capacidad de poder procesar y analizar muchos datos, que puedan llevar a conclusiones científicas irrefutables en base a las cuales se toman decisiones y se hacen recomendaciones. En el ámbito de la salud, las enfermedades se pueden clasificar en endógenas (propias del huésped), exógenas (por acción de agentes externos sobre el huésped) y las ambientales (causadas por condiciones en el medio).

En el primer tipo de enfermedades, las endógenas, caen una gran cantidad de problemas de salud, entre las que se encuentran: predisposición al cáncer, enfermedades autoinmunes, diabetes y otras. Las mismas tienen componentes genéticos muy altos, siendo hereditarios y la capacidad de atacar las razones de la enfermedad, más allá de los síntomas, ha sido relativamente elusiva hasta tiempos



recientes. En este caso, análisis genético de una población muy grande (grandes datos), puede permitir hacer asociaciones en el ADN para encontrar factores relacionados entre los genes y la enfermedad, sin descartar otros factores como los ambientales, efectos de los medicamentos y otros, que pueden habilitar mejores diagnósticos, mucho antes de que aparezcan los síntomas. Del mismo modo, la tecnología de procesamiento de datos puede estar siendo alimentada permanentemente con publicaciones científicas, libros, artículos y datos públicos para poder descubrir nuevos tipos de patrones y tratamientos.

Dentro del grupo de enfermedades exógenas se encuentran la infecciosas, parasitarias (Como el paludismo, también conocido como malaria) y venéreas (Como el VIH), y además de la posibilidad de hacer análisis sintomáticos y de los tratamientos como en el caso de las endógenas, una inteligencia artificial autónoma puede detectar de manera temprana patrones de enfermedad e aumento de incidencias para que, de modo preventivo, se puedan preparar las instituciones de salud para tratar una posible epidemia. Esto es sumamente importante, ya que uno de los factores primordiales para la supervivencia en casos de ciertas enfermedades es la detección temprana y la contención de la misma para evitar que afecte a más individuos.

Para el caso de las enfermedades ambientales, la tecnología puede aplicarse de manera similar a las exógenas, pero adicionalmente se pueden agregar variables correspondientes al medio ambiente, entre las que podrían estar los patrones de estaciones, análisis de niveles de contaminación del aire y agua y otras.

Un caso especial de las enfermedades endógenas se puede encontrar en la malnutrición. Un mito que aún es necesario vencer es el de que comer equivale a alimentarse bien. A pesar de que se coman las tres comidas, es muy posible que gran parte de ese alimento consista de calorías vacías, que aportan pocos nutrientes para el cuerpo y mucha energía, satisfaciendo el hambre, pero no es balanceado para el cuerpo. Para este caso es necesario no sólo información sobre los ciudadanos, sino también sobre la distribución de alimentos y niveles de consumo, para poder detectar y evitar los efectos de la malnutrición en la población. Poder conocer los requerimientos nutricionales reales de cada individuo y facilitarle la adquisición de los alimentos adecuados (ya sea por la distribución de los mismos, la disponibilidad o el costo) puede permitir que los problemas de hambre se resuelvan en un tiempo más corto, más aún si el sistema tiene independencia para tomar decisiones que mejoren este proceso de manera dinámica.

Un punto adicional que engloba el análisis expuesto anteriormente es que, como bien se indicó, el factor primordial son los datos. La Tecnología Usable está poniendo los sensores y las mediciones al alcance de todos los individuos. Sensores que pueden leer constantemente la presión arterial, medir el ejercicio físico que se está haciendo,

sudoración, temperatura corporal, niveles de glucosa en la sangre (entre otros valores) y otros parámetros que cada día crecen más en complejidad, pueden monitorear constantemente el estado de salud de los individuos en una sociedad y evitar los efectos de las enfermedades lo antes posible.

Es importante considerar que para enmarcar el estado actual de la tecnología, muchas de las capacidades mencionadas en este artículo ya existen hasta cierto punto en la filosofía de FOSS (Software Libre y Abierto), principalmente en torno a proyectos dentro de la Fundación Apache (Hadoop, Mahout, entre otros), OpenStack y Java y empresas multinacionales de diferentes tamaños también se encuentran haciendo desarrollos importantes alrededor de esta tecnología.

## **6.- Conclusiones**

Con las últimas tendencias de la tecnología, representadas en la Nube, Internet de las Cosas y Big Data y las mejoras en procesos para tener cada vez mayor capacidad de cómputo en menos espacio, un sistema que parece ser tan complejo como la inteligencia artificial autónoma será completamente portable en apenas unos años, lo que pudiese significar que cualquier elemento pudiera estar siendo procesado por una inteligencia que va más allá de monitorear variables, puede complementar el razonamiento humano, convirtiéndose en una inteligencia aumentada, la natural junto a la artificial.

Los datos, en toda su variabilidad, volumen y velocidad, podrán ser procesados para tomar las decisiones adecuadas prácticamente en tiempo real y pueden servir para el mejoramiento de la condición humana, en muchos ámbitos del día a día.

Políticas de salud definidas a través de entes como el Instituto Nacional de los Seguros Sociales potenciadas por este tipo de tecnología pueden tener un efecto transformativo en el sistema de salud venezolano, guiado también muy de cerca por el aumento del acceso a internet y la tecnología a través de CONATEL, puede llegar a todos los ciudadanos del país.

## 6.- Referencia Bibliográfica

- STERNBERG, R.J.; DETTERMAN, D.K. (1988). ¿Qué es la inteligencia? Enfoque actual de su naturaleza y definición. Madrid: Pirámide.
- SNIJDERS, C.; MATZAT, U.; REIPS, U.-D. (2012). "'Big Data': Big gaps of knowledge in the field of Internet". *International Journal of Internet Science* 7. 1–5.
- KOVAHI, RON; PROVOST, FOSTER (1998). "Glossary of terms". *Machine Learning*. 271–274.
- POOLE, DAVID; MACKWORTH, ALAN; GOEBEL, RANDY (1998). "Computational Intelligence: A Logical Approach". New York: Oxford University Press.
- ANZAI, YUICHIRO (2012) "Pattern Recognition and Machine Learning". Morgan Kaufman.
- GIRARD, JOHN; KLEIN, DEANNA; BERG, KRISTI (2015). "Strategic Data-Based Wisdom in the Big Data Era". IGI Global.
- KIRK, MATTHEW (2014). "Thoughtful Machine Learning" O'Reilly Media, Inc.
- MAY, MIKE (13 de Junio, 2014). "Big Biological Impacts From Big Data": [http://www.sciencemag.org/site/products/lst\\_20140613.xhtml](http://www.sciencemag.org/site/products/lst_20140613.xhtml)
- LIU, DONG (12 de Enero, 2015). "Artificial Intelligence Sector Analysis (Landscape Overview)": <http://insights.venturescanner.com/2015/01/12/artificial-intelligence-sector-analysis-part-i/>
- AMYX, SCOTT (Diciembre, 2014) "Wearing Your Intelligence: How to Apply Artificial Intelligence in Wearables and IoT" <http://www.wired.com/2014/12/wearing-your-intelligence/>