



UBA #1 Iberoamérica ranking QS



IALAB

www.ialab.com.ar



Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura



Cátedra UNESCO en Sociedades del
Conocimiento y Gobernanza Digital
Universidad Nacional del Sur
Bahía Blanca, Argentina

Seis niveles de automatización

Matizar la relación personas-máquinas
y máquina-máquina en el ecosistema digital

Autores:

Juan Gustavo **Corvalán**

Elsa **Estevez**

Enzo **Le Fevre**

[Ver infografía](#)



SEIS NIVELES DE AUTOMATIZACIÓN

Matizar la relación personas-máquinas y máquina-máquina en el ecosistema digital¹

Juan G. Corvalán², Elsa Estevez³ y Enzo Le Fevre Cervini⁴

La autonomía que se atribuye a los sistemas inteligentes es objeto de debate. El diccionario de Cambridge define a la autonomía como el poder o el derecho de gobernarse a sí mismo [1] . Por su parte, el informe de la Comisión de Asuntos Jurídicos con recomendaciones a la Comisión Europea para la creación de una directiva relativa a las normas de legislación civil en materia de robótica del 17/1/2017 establece, que la autonomía de un robot puede definirse como la capacidad de tomar decisiones y aplicarlas en el mundo exterior, con independencia de todo control o influencia externos; que esa autonomía es puramente tecnológica y que será mayor cuanto mayor sea el grado de sofisticación con que se haya diseñado el robot para interactuar con su entorno [2].

Como se evidencia, no hay consenso unánime sobre qué implica que un sistema sea autónomo. Tampoco se pretende que lo haya, porque la autonomía no es una cualidad unidimensional ni binaria, sino es una dimensión continua que presenta matices, por lo que no es tan sencillo determinar cuándo un comportamiento realizado por una máquina va a ser autónomo y cuándo no [3] . Además, esta depende en gran medida del contexto en el que se inserte el sistema del que se trate. Veamos un ejemplo.

En 2018 Google lanzó un servicio en la nube que permite construir y entrenar automáticamente a un algoritmo de aprendizaje profundo [4]. Así el jefe científico de Google Cloud afirmó: “Automatizar el desarrollo

[1] Equipo de investigación y diseño: María Victoria Carro, Giselle Heleg y María Victoria Mafud

[2] Director del Laboratorio de Innovación e Inteligencia Artificial, Facultad de Derecho, Universidad de Buenos Aires

[3] PhD en Ciencias de la Computación

[4] Project Manager, Unión Europea

de este tipo de sistemas podría conseguir que la IA esté al alcance de muchas más personas” [5]. Se trata, nada más ni nada menos, que de una inteligencia artificial capaz de crear a otra inteligencia artificial [6].

Si tenemos en cuenta la completa ausencia de intervención humana en este ejemplo, tal vez podríamos decir que se trata del nivel más elevado de autonomía. La actuación humana, la interacción hombre-máquina o cualquier otra forma de referirse al rol que juega el actuar humano en la actividad de los robots, es un factor que toman en cuenta la gran mayoría de conceptualizaciones de autonomía, como un elemento definitorio.

En atención a esta independencia de control se han postulado tres modelos de sistemas según el grado de interacción hombre-máquina [7] : (1) *Man in the loop*, cuando la IA necesita aportes humanos a intervalos de tiempo regulares para poder llevar a cabo sus acciones; (2) *Man on the loop*, si la máquina es capaz de actuar por sí misma a partir de una programación previa, pero el humano puede intervenir interrumpiendo o modificando las acciones del robot en cualquier momento; y (3) *Man out of the loop*, un modelo en el que la máquina actúa de manera independiente durante ciertos períodos de tiempo y, en estos intervalos, el ser humano no tiene influencia sobre las acciones del robot. Como puede verse, se ha puesto mucho énfasis en la duración de los intervalos en los que existe intervención o control humano y en los que no.

Esta es una de las muchas escalas de matices que pueden proponerse sobre la autonomía de sistemas. Incluso, existe otra ampliamente aceptada en materia de vehículos autónomos. Nos referimos al famoso estándar J3016 del año 2014 de la Sociedad de Ingenieros de Automoción (SAE) [8]. Las categorías que conforman este estándar son las siguientes:

0- Control humano. Todas las funciones operativas son controladas por el conductor humano. El conductor realiza a tiempo completo todas las tareas asociadas a la conducción, incluso cuando son mejoradas mediante sistemas de alerta o intervención. El conductor sigue teniendo pleno control del vehículo y se necesita su presencia constante.

1- Asistencia al conductor. En este nivel, tecnologías tales como los frenos

automáticos de emergencia (AEB) o el control de crucero, pueden ser controladas por el vehículo para asistir al conductor. Es decir, que el sistema de ayuda a la conducción, desarrolla una tarea específica, como la conducción dinámica lateral o longitudinal utilizando la información del entorno del vehículo, mientras que el conductor realiza el resto de tareas de conducción.

2- Automatización parcial. Al menos dos funciones automatizadas simultáneas pueden ser controladas por el vehículo. Por ejemplo, la conducción longitudinal y lateral simultáneamente o la aceleración y la dirección pueden asistir al conductor, si bien este siempre tiene el control del vehículo. El vehículo logra determinar autónomamente peligros y puede, de manera autónoma, interrumpir la automatización si el conductor no responde a la solicitud de control y/o puede parar el vehículo de forma autónoma para evitar peligros (como en caso de detectar la llegada de otro vehículo cuando está yendo marcha atrás).

3- Autonomía condicional. En ciertas circunstancias, el coche puede gestionar todas las funciones críticas relacionadas con la seguridad, pero se espera que el conductor tome el control cuando sea advertido. El conductor debe permanecer alerta incluso cuando el coche esté bajo control autónomo. Es decir, que el sistema de conducción automatizada desarrolla todas las tareas de la conducción con la expectativa de que el conductor responda adecuadamente a la petición de intervención por parte de éste. Si bien no es necesario que un humano realice la supervisión de la conducción automatizada, siempre debe estar en una posición adecuada para reanudar el control.

4- Conducción altamente automatizada. El vehículo es totalmente autónomo en algunos escenarios de conducción, si bien no en todos. La autonomía total en entornos predefinidos requiere que no haya intervención de un conductor humano. Es decir, que aquí también el sistema de conducción automatizada realizará todas las tareas de conducción; pero a diferencia del nivel anterior, lo hará incluso si el conductor no responde adecuadamente a la petición de intervención por parte de este.

5- **Autonomía total o plena.** El coche es totalmente capaz de circular de forma autónoma en cada situación y en cada entorno sin necesidad de intervención de un conductor humano.

Ahora bien, cabe mencionar que también se han elaborado otras escalas para sistematizar los niveles de automatización de vehículos. La Administración Nacional de Seguridad del Tráfico en las Carreteras (NHTSA) de Estados Unidos emitió en 2013 una "Declaración preliminar de política relativa a la automatización de Vehículos" donde consigna 5 niveles: 0 (no automático), 1 (una función específica automática), 2 (funciones combinadas automáticas), 3 (automatización de conducción autónoma limitada) y 4 (automatización de conducción autónoma completa). Sin embargo, finalmente primó en la industria el estándar confeccionado por la SAE; de hecho la NHTSA ha adoptado este sistema de clasificación en septiembre de 2016, abandonando el suyo propio [9].

Muchas legislaciones de distintas partes del mundo adoptan el estándar J3016 de SAE en sus regulaciones sobre vehículos autónomos. Comenzando por algunos estados de Estados Unidos, la ley de Maine establece que un sistema de conducción automatizada se utiliza específicamente para describir los niveles 3, 4 y 5 de automatización.

En el caso de Australia, en el año 2016, los ministros de transporte acordaron un programa de reforma gradual para que los vehículos correspondientes al nivel de 3 de la clasificación J3016 de SAE puedan operar de manera segura y legal en las carreteras australianas antes de 2020, y los vehículos altamente y completamente automatizados a partir del mismo año. Luego, en 2017 se publicaron las pautas para los ensayos de vehículos automatizados en Australia y el documento de política adjunto tras la aprobación por parte de los ministros de transporte.

Canadá ha sido otro país en seguir el estándar. La Comisión Permanente de Transportes y Comunicaciones del Senado ha dictado un documento en el año 2018 [10] , que adopta la clasificación de J3016 de SAE. Finalmente en Japón, la legislación Public-Private ITS: Initiative/Roadmaps 2017 adhiere a la clasificación J3016 de SAE y denomina a los vehículos correspondientes al nivel 3 "vehículos altamente automatizados" y los

vehículos del nivel 4 y 5 "vehículos completamente automatizados".

Tomando como modelo estas propuestas sobre matices de autonomía, desde IALAB intentamos elaborar una escala en materia de procesamiento del lenguaje natural. A continuación, se explican los niveles que hemos elaborado y se ilustran con un caso de una tarea de revisión de un texto realizada por una persona a través de un sistema tipo "Word".

En un **primer nivel de automatización**, el procesador de texto permite mejorar la forma de copiar y pegar un posible sinónimo, aunque es la persona quién tiene que llevar adelante la tarea. Esta fue la lógica que imperó en las primeras generaciones de procesadores de texto. Se trata de una "máquina de escribir" muy mejorada. El funcionamiento del sistema depende de la activación humana, a partir de un menú cerrado de opciones (al menos hasta que se actualiza el programa). En este primer nivel de automatización, la intermediación de la máquina no altera en absoluto la voluntad y el ejercicio del dominio de la persona humana.

Un **segundo nivel de automatización** se vincula con la inserción de una lista de sinónimos dentro del sistema, confeccionada en primer lugar por un humano, quien determinó qué opciones de sinónimos se ofrecerían para cada palabra y posteriormente cargada al sistema por el programador, para que la persona pueda posarse en la palabra y que luego aparezca un catálogo taxativo e inmutable de posibilidades. Aquí se profundiza el nivel de automatización a partir de mejorar la "asistencia digital". El sistema comienza a concentrar tareas que eran realizadas por las personas. Por ejemplo, que el programa informático subraye posibles textos mal escritos de manera autónoma.

Un **tercer nivel de automatización** puede ejemplificarse del siguiente modo. Aunque el sistema ofrece una lista de sinónimos, esta no se encuentra preconfigurada, sino que cuando la persona decide reemplazar la palabra repetida, el procesador de textos se sirve de internet para buscar sinónimos en la web, de la cual extrae palabras para confeccionar la lista de sugerencias. La selección de las palabras puede

provenir de una lógica de reglas que el programador fijó o bien, a partir de conectar con un sitio concreto del cual el sistema extraerá el sinónimo según la palabra que se trate (por ejemplo, de una página web de sinónimos que se actualizan periódicamente). Incluso, se podría ir más allá, y el propio sistema podría establecer un ranking de palabras más usadas como sinónimos, y luego seleccionar la que tenga mayor peso estadístico.

Cada vez que el usuario activa el sistema, podrían aparecer otras palabras u otras sugerencias, de modo tal que podemos hablar de un cierto nivel básico de autonomía focalizada en la búsqueda y selección de opciones, aunque luego se subordina a la acción del usuario humano que utiliza el sistema. Hasta acá, podemos advertir que en el segundo y en el tercer nivel de automatización, se configuran dos grandes fenómenos que coexisten. Por un lado, se robustece la asistencia digital para tomar la decisión acerca de cuál será el texto que finalmente corresponde reemplazar. Es decir, se incrementa y multiplica la cantidad de opciones sugeridas, lo que debería ampliar la base de conocimiento para tomar una decisión. Esta automatización de tareas secundarias o micro-tareas asociadas a disposición del usuario/a, mejora el ecosistema digital en torno a las sugerencias. Luego, la persona decidirá actuar o no en función de ellas. Por otra parte, en el tercer nivel se advierte un grado de autonomía asociado a la búsqueda inteligente de posibles opciones, pero ello no cercena o limita la voluntad humana. Es la persona (humano en control) quien decide qué palabras se modifican en el texto.

En síntesis, en los tres primeros niveles de automatización, las personas humanas están siempre en control acerca de las tareas principales. En el ejemplo que venimos trabajando, recordemos que se trata de analizar los niveles de automatización y autonomía en un procesador de texto como Microsoft Word, y ejemplificamos con tareas que consisten en revisar textos y modificarlos en función de errores de gramática, sintaxis o encontrar sinónimos a palabras que se suelen repetir con frecuencia en los párrafos de documentos de trabajo.

En el **cuarto nivel de automatización** se advierte una diferencia cualitativa. A lo anterior, se agrega un componente más robusto de

autonomía informática. Continuemos con nuestro ejemplo. Si se usan las nuevas funcionalidades de Microsoft 365, el usuario puede seleccionar una oración completa o incluso un párrafo, y habilitar una función de autocorrección del programa. Esta corrección abarca los errores gramaticales y sustitución de palabras repetidas. Aunque la persona sigue en control, es posible salir del “modo sugerencia” para automatizar, en sentido estricto, la función de copiar y pegar. El humano activa la consecuencia prevista para la acción de corrección y aunque el sistema advierta errores, no está habilitado o no puede autocorregirlos sin una voluntad humana que lo active.

Ahora bien, una vez que esto último ocurre, el programa presenta una *autonomía completa para esa tarea predefinida*; esto es, la sustitución de palabras repetidas sólo dentro de la oración o párrafo indicado por el funcionario. Si el procesador de texto advierte una palabra repetida, procederá automáticamente a insertar un sinónimo que el propio sistema estime que corresponde en función de los disponibles para esa palabra. Esta acción también podría presentar muchos matices de automatización.

Por ejemplo, se puede programar al procesador de texto para que cuente las veces que se repitió una palabra en todo el documento y también en cada página, para luego elegir el sinónimo que menos se repita y que, además, tenga en cuenta el contexto gramatical del párrafo para optimizar la elección. Como luego veremos con mayor detalle, esta autonomía traslada el problema a la faz de programación del sistema. Si el órgano y el funcionario no participan o controlan esa programación, entonces el algoritmo, en los hechos, se ejecuta por fuera de las normas que habilitan el ejercicio de la competencia.

Siguiendo nuestro ejemplo, si luego de esa elección automatizada que ejecuta el programa, no existe una revisión humana y la tarea se ejecuta.

Aquí comienzan a plantearse los problemas en la era de la IA y la automatización. Por un lado, se habilita de manera deliberada al sistema a que proceda de manera autónoma, a elegir y luego cortar palabras incorrectas o con problemas de gramática y sintaxis para reemplazarlas

de forma autoejecutable. Aquí es crítico determinar si la persona participó en la programación del sistema, bajo una lógica que hoy se conoce como co-diseño y co-desarrollo. Esto es importante cuando se trata de que los sistemas automatizados reflejen o espejen voluntades y criterios de decisión humanos.

Por eso es fundamental que se co-diseñen, auditen y controlen periódicamente las funciones que ejecutan los sistemas que están entre el nivel 4 y 6 de automatización, más allá de que en muchos casos (por ejemplo, para actividades que realiza el Estado) no pueden ser habilitados. Si continuamos en nuestro ejemplo, deberíamos estar en control acerca de las listas de sinónimos o las lógicas por las cuáles el programa elige palabras para que sean autoreemplazadas. Por lo tanto, en este escenario, quien deliberadamente habilita al programa para que autoejecute la tarea de selección y reemplazo, pierde en el control humano.

Ahora bien, aunque elegimos un ejemplo muy rudimentario, lo cierto es que las normas no prevén estos fenómenos de automatización en donde la autonomía de los sistemas automatizados puede quebrar la lógica competencial humana [11]. Entre otras medidas, si queremos evitar que esto se configure, es indispensable que la tarea sólo se despliegue a modo de “primer borrador” o sugerencia. En ciertas actividades, son los humanos quienes deben estar en control para tomar decisiones.

El **quinto nivel de automatización** profundiza la autonomía del programa. Aquí no es necesario que las personas seleccionen previamente la frase o párrafo que debe ser objeto de corrección por parte del sistema, sino que a medida que la persona va confeccionando el documento, el programa de forma autónoma autoejecuta la corrección de errores y reemplaza palabras repetidas inmediatamente después de detectarlos. El programa, siguiendo siempre el mismo ejemplo del procesador de texto tipo Word, es autónomo en las funciones o escenarios de reemplazo de palabras repetidas y corrección de errores gramaticales. Aquí no tiene lugar ni la habilitación de la función por parte del humano, y mucho menos el control en la selección de palabras que deben ser reemplazadas o el sinónimo que ocupará su lugar. Su

intervención como usuario/a es nula, en la tarea específica de corregir aspectos formales del acto.

El **sexto nivel de automatización** involucra uno de los máximos grados de autonomía que puede presentar un sistema inteligente.

Continuando con nuestro ejemplo, Microsoft 365 permite a una persona seleccionar una oración incompleta y pedirle al sistema que la complete. En este caso, el sistema genera texto por sí solo. La versión más autónoma en este paradigma se evidencia en el sistema más avanzado de lenguaje natural: GPT-3 [12] . Este es capaz de generar de forma autónoma cualquier tipo de texto, desde poesías hasta artículos periodísticos, acordes de guitarra y códigos informáticos, sin necesidad de intervención de un operador humano. En este último caso, basta con que el humano le proporcione al sistema un título o una consigna.

Notas al final

[1] Disponible en: <https://dictionary.cambridge.org/es/diccionario/espanol-ingles/autonomia>

[2] Informe del 27/1/2017 de la Comisión de Asuntos Jurídicos con recomendaciones a la Comisión Europea para creación de una directiva relativa a las normas de legislación civil en materia de robótica. Disponible en <http://www.europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A8-2017-0005&format=XML&language=ES>

[3] ANDERSON, T. L., y DONATH, M.: «Animal behavior as a paradigm for developing robot autonomy», en *Robotics and autonomous systems*, vol. 6, n.º 1-2, 1990, pp. 145-168; STEINFELD, A., FONG, T., KABER, D., LEWIS, M., SCHOLTZ, J., SCHULTZ, A., y GOODRICH, M.: «Common metrics for human-robot interaction», en *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, ACM, 2006, pp. 33-40; HASELAGER, W. F.: «Robotics, philosophy and the problems of autonomy», en *Pragmatics & Cognition*, vol. 13, n.º 3, 2005, pp. 515-532

[4] El aprendizaje profundo es una técnica de redes neuronales. Los sistemas de IA que utilizan redes neuronales artificiales (RNA) frecuentemente obtienen mejores resultados para reconocer patrones cuando se trata de analizar cantidades masivas

de datos (Big Data 37). Son más eficientes, requieren intervención humana reducida y pueden trabajar datos que los informáticos llaman “no estructurados” o no organizados bajo criterios concretos. Un Excel con atributos y datos etiquetados es información estructurada. Lo que la gente publica en las redes sociales es un ejemplo de información no estructurada. Las redes neuronales complejas que tienen “capas ocultas”, imitan o copian ciertos rasgos de los procesos neuronales de los cerebros humanos, que procesan la información a partir de neuronas, sinapsis, dendritas y axones. Las técnicas que utilizan los expertos en IA consisten en desarrollar algoritmos que implementan redes neuronales para reconocer la regularidad de los datos o patrones. Sin embargo, esto no significa que las redes neuronales artificiales funcionen igual a las biológicas. Ampliar en CORVALÁN JUAN G. “Inteligencia artificial, automatización y predicciones en el derecho” en DUPUY DANIELA, CORVALÁN JUAN G. “Cibercrimen III” Edictorial BdeF, 2020

[5] KNIGHT WILL “Construya una inteligencia artificial a medida con lo nuevo de Google” MIT Technology Review, 2018. Disponible en: <https://www.technologyreview.es/s/9941/construya-una-inteligencia-artificial-medida-con-lo-nuevo-de-google>

[6] Otro caso similar podría ser el Alpha Go Zero, una IA capaz de auto-entrenarse y vencer a cualquier ser humano en el juego Go. Sobre esta IA véase: CORVALÁN JUAN G. “Presentación. Inteligencia Artificial. Automatización y predicciones en el derecho” en DUPUY DANIELA, CORVALÁN JUAN G, KIEFER MARIANA “Cibercrimen III” Editorial B de F pp. 33

[7] HARBERS M., PEETERS M.M.M., NEERINCX M.A. “Perceived Autonomy of Robots: Effects of Appearance and Context” en ALDINHAS FERREIRA, M. I., SILVA SEQUEIRA, J., TOKHI, M. O., KADAR, E., Y VIRK, G. S. (EDS.), A World with Robots, Springer, Cham, 2017, pp. 19-33. MIRO LLINARES FERNANDO “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por Robots” Revista de Derecho Penal y Criminología, 3.ª Época, n.º 20 (julio de 2018), págs. 87-130

[8] El estándar J3016 puede verse en: <https://www.sae.org/news/press-room/2014/10/sae-international-technical-standard-provides-terminology-for-motor-vehicle-automated-driving-systems> Asimismo, se ha publicado una tabla complementaria que resume los niveles de automatización para vehículos de carretera. Véase: https://www.sae.org/binaries/content/assets/cm/content/news/pressreleases/pathway-to-autonomy/automated_driving.pdf

[9] Otras clasificaciones han sido la del Instituto Federal de Investigación de Carreteras (BASt) en Alemania en 2013 que recoge 5 niveles; y la clasificación de la Organización Internacional de Constructores de Automóviles (OICA), la cual comenzó empleando la clasificación del BASt, pero la ha adaptado introduciendo el nivel 3 de automatización condicionada, y pasando a considerar también 6 niveles de alineándose un poco más con la clasificación de la SAE y aceptando múltiples conceptos que esta introduce. Véase IBAÑEZ “De 0 a 5: cuáles son los diferentes niveles de conducción autónoma, a fondo” Xataka, 2018. Disponible en: <https://www.xataka.com/automovil/de-0-a-5-cuales-son-los-diferentes-niveles-de-conduccion-autonoma>

[10] “Driving Change. Technology and the future of the automated Vehicle” Disponible en: https://sencanada.ca/content/sen/committee/421/TRCM/Reports/COM_RPT_TRCM_AutomatedVehicles_e.pdf

[11] La autonomía de una máquina es un concepto artificial que refleja dos fenómenos. Por un lado, que este sistema refleja la voluntad de un programador, que a su vez podría reflejar la voluntad de quien toma las decisiones en cuanto a cuáles serán las reglas decisionales dentro del programa informático. Por otra parte, también es posible que adquiera autonomía (además de la elaborada de los programadores), porque está basado en la aplicación de un sistema de inteligencia artificial que aplica aprendizaje profundo. Mucho más aún, si se trata de uno basado en un sistema de aprendizaje no supervisado. Ver CORVALÁN JUAN G. DANESI CECILIA “Responsabilidad civil de la inteligencia artificial” Colaboradora: María Victoria Carro, apart. 3 “De “Espejito, espejito...” a Skynet. La autonomía de la inteligencia artificial” Tomo III Tratado de Inteligencia Artificial y Derecho.

[12] Sobre GPT-3 ver CORVALÁN JUAN G. “Los límites de la inteligencia artificial. Correlaciones, causalidad, Shakira, GPT-3 y Alicia en el país de las maravillas.” Colaboradora: María Victoria Carro, Tomo I Tratado Inteligencia Artificial y Derecho.

