

- ▶ **Entrevista a Flavia Costa: perspectivas sobre la inteligencia artificial en el contexto social y tecnológico.**

**Interview with Flavia Costa: perspectives on artificial intelligence in the social and technological context.**

**Maximiliano Peret y Mariana Ferrarelli**

Cómo citar: Peret, M. y Ferrarelli, M. (2024). Entrevista a Flavia Costa: Perspectivas sobre la Inteligencia Artificial en el contexto social y tecnológico. *Revista Argentina de Comunicación*, 12(15), 92-100.



Flavia Costa es una de las voces más influyentes a la hora de hablar de la Inteligencia Artificial y su influencia en nuestra sociedad. A lo largo de una carrera dedicada a la investigación y el análisis crítico, ha trabajado sobre cómo las tecnologías avanzadas moldean nuestras vías en múltiples dimensiones. Su trabajo se centra en analizar los vínculos entre la tecnología, la cultura y la sociedad con especial énfasis en la necesidad de una perspectiva crítica.

Uno de sus aportes más significativos es el concepto de Tecnoceno, a través del cual Costa propone pensar que hemos ingresado en una nueva etapa histórica donde la tecnología se convierte en un factor dominante en la organización de la vida y el entorno. Su análisis va más allá de la Inteligencia Artificial como herramienta y la posiciona como una metatecnología, un sistema que estructura y transforma otros sistemas en áreas como la logística, la educación, la salud y la comunicación. Este enfoque le ha permitido desarrollar una mirada interdisciplinaria que conecta los estudios sociales, filosóficos y técnicos invitándonos a reflexionar sobre cómo y por qué adoptamos ciertos avances tecnológicos y a qué costo.

Costa también es responsable de proyectos de investigación centrados en las formas de vida infotecnológicas, en los cuales indaga cómo la digitalización y el control informático afectan nuestras experiencias cotidianas. En su trabajo en el Tecnocenolab ha abordado temas como la ética en el diseño de sistemas de IA y los desafíos de las inteligencias artificiales generativas para la democracia a la vez que plantea una visión crítica sobre el papel de estas tecnologías en la sociedad actual.

En esta entrevista indagamos en su visión sobre el estado actual de la inteligencia artificial y sus implicancias éticas y políticas además de pensar en los desafíos futuros que este campo plantea para el desarrollo de una sociedad justa y equitativa.

**¿Cómo llegás al estudio de la IA? Conocemos tu trayectoria científica, ¿podrías mencionar dos o tres hitos destacados que te llevan a interesarte por este fenómeno?**

En principio diría que es un tema obligado de mi campo de estudios. Pertenezco desde 1995 a un espacio de trabajo en la Facultad de Ciencias Sociales de la UBA que es la cátedra de Informática y sociedad, hoy rebautizada Técnica, cultura y sociedad, fundada originalmente por Héctor “Toto” Schmucler. Se trata de un espacio particularmente marcado por la perspectiva crítica, en el cruce entre filosofía de la técnica, ciencias de la comunicación y estudios mediales, y mi ingreso allí convergió con el desembarco de internet comercial en la Argentina. En esos años una de nuestras premisas era construir conocimiento sobre ese acontecimiento desde el Sur: ¿qué significaba la “red de redes” como proyecto comunicacional, cultural y político? Sobre el final de esa década e inicios del siglo XXI se suma a esa pregunta la cuestión ‘bio’, que responde en parte a la conciencia de que las tecnologías infocomunicacionales y la *episteme* de la información en que se fundamentan ya estaban abriendo un nuevo ciclo. Me concentré en la pregunta por el cruce entre las políticas y las tecnologías de la vida. En 2003 empecé a estudiar lo “bio”: el bioarte, la biopolítica, la biología molecular. Con Pablo Manolo Rodríguez hicimos una investigación interdisciplinaria sobre las nociones de vivo y de viviente en el arte, el derecho, la biología sintética, la comunicación, la filosofía. El cruce entre políticas y tecnologías de lo viviente fue uno de los ejes de mi Tesis doctoral.

En 2010 empecé a profundizar en las formas de vida infotecnológica, o cómo son nuestras formas de vida atravesadas por la doble tendencia a la tecnificación y a la politización de vida que caracteriza nuestro tiempo. En 2016 me interesé en los estudios sobre vigilancia (el modo en que las redes informáticas permiten una “trazabilidad” permanente) y eso implica pensar en la infraestructura de las redes como parte de una tecnología política.

Finalmente otro hito es la pandemia, cuando para muchos quedó en evidencia el salto de escala: la pandemia del covid reveló a buena parte de la población mundial que estaba siendo parte de un acontecimiento de escala planetaria, que es posible como parte de la historia del Tecnoceno. Y entre diciembre de 2022 y marzo de 2023 se produce un giro dentro de esta zona: a partir de la puesta en disponibilidad pública del chat-gpt, un chat de IA generativa, y tras la carta de Future of Life que pide una “pausa” en el desarrollo de IA, se desata un “furor normativo” que todavía está en marcha.

**En función de este recorrido que mencionás, ¿por qué catalogás a la IA como metatecnología? ¿Qué supuestos subyacen a esta definición?**

En 2023, en Tecnocenolab hicimos una investigación sobre los desafíos de las IA generativas (y otras IA) para la calidad democrática. Para abordar ese problema definimos una perspectiva analítica desde la cual mirar las IA emergentes a partir de cinco claves. Una de ellas es considerar que las IA son metatecnologías. La segunda, que no son herramientas, sino que, junto con los ecosistemas digitales a los que pertenecen, conforman un mundoambiente. Luego, que en determinados usos críticos para la población pueden ser tecnologías de riesgo, y que para abordarlas es necesario, no solo una “ética” de las IA sino también un pensamiento sistémico acerca de las IA, y esto en relación con los valores de las sociedades en los que éstas se insertan. Finalmente, cuando hablamos de estos modelos ellos no son sólo Inteligencia Artificial, sino también Sociedad Artificial. Son sistemas que analizan, automatizan, gestionan lo social y por lo tanto, hay saberes de las ciencias sociales específicos sobre los problemas y los riesgos relativos a ciertos usos de las IA: desinformación planificada, vigilancia masiva, sesgos, uso malicioso, calificación social. En las ciencias sociales sería valioso fortalecer la formación en saberes, métodos y tareas propias de esos equipos técnicos, para incorporar la mirada analítica y crítica.

Ahora bien, en este marco, metatecnologías significa que las IA son tecnologías que actúan sobre muy distintos conjuntos técnicos. Operan sobre otros sistemas y subsistemas: de logística, de energía, de salud, de educación, de organización y gestión de lo público, de comunicación. Son metatecnologías porque apuntalan otras tecnologías; automatizan tareas y aceleran el desempeño de otros conjuntos técnicos.

**¿En qué sentidos considerarás que esta etapa de la IA representa un riesgo para los humanos? ¿Tenemos que tomarnos en serio la carta que sacaron los popes de la industria en 2023 cuando hablaban de eso?**

Una de las formas de encarar esta pregunta es decir que las IA representan riesgos según su escala. Una IA que permite optimizar un láser para metrología no implica un gran riesgo; una IA que puede ayudar a entrenar a una persona con escasos saberes de química para que alcance conocimientos similares al nivel de doctorado, ya es otro nivel de riesgo. Por otro lado, la controversia por el así llamado “riesgo existencial” no es un invento mediático ni de los grandes magnates de la tecnología: es una controversia que se está desarrollando en los más altos niveles de la investigación. Están metidos en ella los tres Premio Turing 2018: Yoshua Bengio, Yan LeCun y Geoffrey Hinton (flamante Premio Nobel de Física 2024).

Que las IA pueden ser tecnologías de riesgo en relación con los usos que se haga de ellas es la perspectiva de la ley de la Unión Europea, que sigue en eso la postura del filósofo de la información Luciano Floridi. Por mi parte creo que gestionar y analizar los riesgos es parte de la agenda del Tecnoceno. Nosotros llegamos a esto siguiendo el rastro de la pregunta por los accidentes de toda tecnología, y por la preocupación por los “accidentes normales”, que revivió con la pandemia.

## ¿Qué papel deberían jugar los Estados en la regulación del desarrollo y el uso de inteligencia artificial en el ámbito de la comunicación?

Los distintos estados y bloques regionales están avanzando en modelos de gobernanza de la IA, donde suele haber –y creo que es fundamental– alguna ecuación entre desarrollo y seguridad. En ninguna política sería se elige una sola de esas dos instancias; siempre se busca algún punto de equilibrio. Hay modelos como el de la Unión Europea, donde la seguridad está abordada en una ley integral y el desarrollo de IA debe ajustarse a esa ley; y hay otros modelos, como el estadounidense, al menos tal como aparece en el Acta de 2019 que crea la Iniciativa Nacional de IA, donde se deja que el desarrollo se atenga a la legislación ya existente pero se le encarga al NIST, al Instituto Nacional de Estándares y Tecnología, el desarrollo de un marco de riesgo muy detallado para la IA. Es decir: nunca se descarta la seguridad, y aquí uso seguridad en el sentido de *safety*; no de *security* (que es la seguridad en sentido policial, la del delito). Me refiero a que las IA sean seguras para los valores democráticos, que no atenten contra las minorías, ni contra la libertad de expresión, o contra la libertad y la necesidad de informarse. Dicho eso, no tengo ahora mismo presentes normas o recomendaciones ya suscriptas o implementadas a nivel estatal sobre el campo de las comunicaciones. Sí recuerdo, sin embargo, dos documentos: el reciente marco de gestión del riesgo de IA para los derechos humanos que emitió en julio de 2024 el Departamento de Estado de los EE.UU., donde se afirma que “un amplio espectro de riesgos asociados con la IA afectan el disfrute y ejercicio de los derechos humanos, incluidos la privacidad, la igualdad ante la ley, la libertad de opinión y expresión”. Y donde sugiere acciones para enfrentar esos riesgos. Y también el informe “La IA como bien público: garantizar el control democrático de la IA en el espacio de la información”, que publicó en febrero de este mismo año el Forum sobre Información y Democracia.

**¿Qué principios éticos considerarás fundamentales para guiar el diseño y la implementación de estas tecnologías que puedan afectar al espacio de la información y la comunicación?**

Les comento lo que está circulando ahora, ya que necesitamos seguir estudiando más. Lo que se recomienda es promover el estudio interdisciplinar de los sistemas de IA y de los usos de la IA que afecten ámbitos críticos para la población, como el acceso libre al espacio infocomunicacional, la libertad de expresión, la privacidad de los datos personales. Lo que hay que hacer también es delimitar los distintos tipos de riesgos (por ejemplo, en los sistemas de curaduría de contenidos, en los sistemas vinculados a la publicidad, en la generación de textos e imágenes, limitar la exposición a la desinformación planificada o a la vigilancia masiva), identificar los incidentes que ocurren en este campo e investigarlos cuidadosamente, con enfoque sistémico, para tener más información sobre cómo están contruidos y cómo funcionan. Los requisitos que se suelen invocar son los de exigir supervisión humana, adherencia a los hechos, moderación de contenido, mecanismos de filtrado que los usuarios pueden aplicar para analizar y marcar contenido potencialmente problemático... Hay mucho por hacer todavía.

**¿Por dónde se empieza a leer una entidad no humana de estas características? ¿Es una máquina, un ensamblaje...?**

Hay diferentes formas de enfocar estos ensamblajes que van más allá de la clásica figura de la máquina-herramienta. Si uno recuerda *Técnica y civilización* de Lewis Mumford, un texto de 1934, él ahí ya hablaba de la “megamáquina” humana que había construido las pirámides de Egipto. Es decir, ya tenía en mente lo que hoy llamamos ensamblajes sociotécnicos, o sistemas sociotécnicos complejos. Donde desde las redes comunitarias a distancia hasta el Estado mismo son “máquinas” con la que interactuamos, que no sabemos bien cómo funciona; una “máquina” parte de cuya productividad es su propia opacidad. Nosotros ya tenemos relación

frecuente con entidades no humanas que son obra humana, por ejemplo las instituciones. La familia misma ya es una “máquina” en la que nosotros somos componentes; actuamos a partir de la familia: no tiene la materialidad física que tiene una computadora, pero sí tiene sus materialidades, sus rituales, su lenguaje, sus olores, sus procedimientos. Otra “máquina” muy curiosa –y sé que esto puede resultar controversial, porque el lenguaje aparece como la superficie de contacto entre lo que llamábamos la naturaleza y la cultura--; otra “máquina”, decía, enormemente rica, es el propio lenguaje. Creo que si creemos en la poesía, si suponemos que la poesía existe y existirá, es porque hay una “máquina” llamada castellano, otra japonés, otra italiano y así, que hace cosas rarísimas a través de “componentes” que van rotando y que son los seres humanos. Y si confiamos en que seguirá existiendo la poesía en esas lenguas es porque intuimos que aún no sabemos todo lo que esas máquinas son capaces de hacer: la cantera parece infinita. Insisto: nosotros interactuamos todo el tiempo con máquinas que están hechas por los humanos y que no las conocemos del todo. Establecemos vínculos con estas entidades no humanas que no son tampoco animadas, como los animales o como las plantas.

### **¿Cómo nos vinculamos con estas máquinas?**

La perspectiva del riesgo es muy defensiva, pero tiene sentido empezar por ahí si estamos frente a sistemas sociotécnicos complejos que automatizan el análisis y la gestión de lo social: si “automatizan” lo social. Si una agencia de gobierno usa IA para automatizar el proceso que lleva a la decisión acerca de si darle o retirarle la ayuda social a una familia, o para otorgar (o no) una vacante para la escuela, o para dictaminar si una persona debe o no ir a la cárcel por un presunto delito, es necesario monitorear y auditar el proceso. Puede ser que en ciertos casos sea necesario automatizar procedimientos, porque ya somos más de 8000 millones de personas en el mundo, y en esto, el juego entre las escalas (la dimensión espacial) y la aceleración de procesos (la dimensión temporal) son dos grandes claves analíticas. Un aparato “manual” de gobierno para una escala de 8000



millones de personas posiblemente sea muy lento. Y nos guste o no, esa “megamáquina” social de la que sabemos tan poco ya se activó. Bueno, de acuerdo: ¿sabemos cuáles son las prioridades? Por un principio básico de prudencia, ¿conocemos los riesgos de usar IA en áreas críticas para la población, como el acceso a la justicia, a la salud, a la educación? ¿Y si en esos casos fuera mejor esperar un poco? ¿Y si evaluamos si no conviene, en algunos casos, utilizar el poder-no, la potencia de no? Es decir: la potencia de no hacer aquello que, técnicamente, puedo hacer pero prefiero prudencialmente esperar y ver primero los así llamado “impactos”, En este sentido diría que hoy, para las ciencias sociales, es momento de dejar de ser observadoras o meramente usuarias de los ecosistemas digitales + IA y es tiempo de pasar a la acción, de ingresar en ellos, de trabajar en su propio terreno. Porque estos problemas son problemas prácticos, políticos. Hay que animarse a hacer ese salto. *Personalmente prefiero hoy equivocarme en lo que estoy pensando acerca de este acontecimiento que no estar pensándolo.*