

***Interacciones. Revista de la Facultad de Ciencias Sociales, Educación y
Comunicación***

ISSN 2796-9010

Núm. 6, año 2026, [pp. 1-20]

**NPC: agentes no humanos y prototipado social como metodología para las ciencias
sociales**

ARK CAICYT : <https://id.caicyt.gov.ar/ark:/s27969010/fjymxw7f3>

Darío Julián Tagnin¹

Resumen

El artículo propone el análisis de personajes no jugables (NPC) como Agentes No Humanos de Comunicación (ANHC) y el prototipado social como una metodología útil para la investigación en ciencias sociales. Desde una perspectiva posfenomenológica, sostiene que estos agentes, dotados de capacidades adaptativas y potencias cognitivas, ya no constituyen simples representaciones computacionales, sino entidades capaces de asumir roles sociales, interactuar entre sí y generar dinámicas emergentes que permiten explorar fenómenos complejos en entornos controlados. Se revisan antecedentes del modelado multiagencial y se analizan desarrollos recientes que emplean agentes generativos para estudiar redes sociales, mercados financieros y comunidades virtuales, destacando su potencial para reproducir procesos de difusión de información, cooperación, conflicto y formación de normas.

¹ Doctor en Epistemología e Historia de la Ciencia (UNTREF), magíster en Comunicación, licenciado en Periodismo y profesor en Comunicación. Se desempeña como docente en la Universidad Nacional de Lomas de Zamora y en la Universidad Nacional de José C. Paz. Sus investigaciones se centran en los agentes no humanos de comunicación y sus implicancias epistemológicas, éticas y sociales, desde la filosofía de la técnica y la posfenomenología. Correo electrónico: jtagnin@unpaz.edu.ar

Fecha de recepción: 15/08/2026. Fecha de aceptación: 1/10/2026

Asimismo, se examina la creciente antropomorfización de estos agentes y la constitución de vínculos afectivos entre humanos y sistemas artificiales, aspecto que complejiza la distinción entre ambos tipos de agencia. Finalmente, el trabajo responde a las principales objeciones dirigidas al modelado social computacional (reduccionismo, implicancias político-epistémicas y cuestionamientos éticos) al argumentar que el prototipado social no busca sustituir la investigación empírica, sino ampliar sus capacidades heurísticas mediante la experimentación con escenarios híbridos donde la agencia se distribuye entre actores humanos y no humanos.

Palabras clave: agentes no humanos de comunicación, NPC, simulación social, modelado multiagencial, prototipado social, posfenomenología

NPCs: Non-Human Agents and Social Prototyping as a Methodology for the Social Sciences

Abstract

This article proposes the analysis of non-player characters (NPCs) as Non-Human Communication Agents (NHCAs) and the use of social simulation as an innovative methodological framework for social science research. Drawing on a postphenomenological perspective, it argues that these agents, equipped with adaptive capabilities and language models, should no longer be understood as mere computational representations but as entities capable of assuming social roles, interacting with one another, and generating emergent dynamics that enable the exploration of complex social phenomena in controlled environments. The paper reviews the foundations of multi-agent modeling and discusses recent developments employing generative agents to study social networks, financial markets, and virtual communities, highlighting their potential to reproduce processes of information diffusion, cooperation, conflict, and norm formation. It also examines the growing anthropomorphization of these agents and the emergence of affective relationships between humans and artificial systems, a phenomenon that increasingly blurs the distinction between human and non-human agency. Finally, the article addresses the main criticisms of computational social modeling (including reductionism, political-epistemic concerns, and ethical objections) arguing that social prototyping is not intended to replace empirical research but to expand its heuristic potential by experimentally exploring hybrid scenarios in which agency is distributed across both human and non-human actors.

Keywords: *Non-Human Communication Agents, NPCs, social simulation, agent-based modeling, social prototyping, postphenomenology*

NPCs: agentes não humanos e prototipagem social como metodologia para as ciências sociais

Resumo

O artigo propõe a análise de personagens não jogáveis (NPCs) como Agentes Não Humanos de Comunicação (ANHHC) e da prototipagem social como uma metodologia útil para a pesquisa nas ciências sociais. A partir de uma perspectiva pós-fenomenológica, sustenta-se que esses agentes, dotados de capacidades adaptativas e potências cognitivas, já não constituem simples representações computacionais, mas entidades capazes de assumir papéis sociais, interagir entre si e gerar dinâmicas emergentes que permitem explorar fenômenos complexos em ambientes controlados. São revisados antecedentes da modelagem multiagente e analisados desenvolvimentos recentes que empregam agentes generativos para estudar redes sociais, mercados financeiros e comunidades virtuais, destacando seu potencial para reproduzir processos de difusão de informação, cooperação, conflito e formação de normas. Do mesmo modo, examinam-se a crescente antropomorfização desses agentes e a constituição de vínculos afetivos entre humanos e sistemas artificiais, aspecto que complexifica a distinção entre esses dois tipos de agência. Por fim, o trabalho responde às principais objeções dirigidas à modelagem social computacional — reducionismo, implicações político-epistêmicas e questionamentos éticos —, argumentando que a prototipagem social não busca substituir a pesquisa empírica, mas ampliar suas capacidades heurísticas por meio da experimentação com cenários híbridos nos quais a agência se distribui entre atores humanos e não humanos.

Palavras-chave: *agentes não humanos de comunicação, NPCs, simulação social, modelagem multiagente, prototipagem social, pós-fenomenologia*

Introducción

El uso de agentes computacionales en simulaciones sociales plantea un problema epistemológico: ¿qué permite conocer una interacción entre entidades técnicas y bajo qué condiciones puede contribuir a investigar lo social? Para abordarlo, tendremos que distinguir categorías que designan dimensiones diferentes. En mi línea de trabajo, los agentes no humanos se refieren a entidades cuya capacidad de intervenir y producir diferencias en un entorno no

requiere atribuirles propiedades humanas; dentro de este conjunto, la categoría de Agentes No Humanos de Comunicación (ANHC) delimita aquellas que participan en la producción, circulación y modulación de intercambios comunicativos (Tagnin, 2024a). NPC (*non-playable character*), por su parte, designa un personaje de un videojuego o entorno simulado que no está controlado directamente por un jugador: identifica una función dentro de ese entorno, compatible con diversas arquitecturas técnicas y grados de autonomía. Los agentes generativos, en el sentido de Park *et al.* (2023), incorporan modelos de lenguaje articulados con mecanismos de memoria, reflexión y planificación para producir comportamientos contextualizados.

Estas categorías corresponden a distintos planos de análisis: ANHC delimita una forma de participación en relaciones comunicativas, NPC identifica una función dentro de un entorno lúdico o simulado, y agente generativo caracteriza una arquitectura técnica. Por ello, pueden concurrir en una misma entidad sin resultar equivalentes: un NPC puede implementarse mediante modelos generativos, mientras que numerosos agentes generativos operan sin asumir el papel de personajes. El análisis prescinde de la noción de inteligencia agéntica, dado que el reconocimiento de agencia no depende aquí de una atribución de inteligencia, de hecho, el análisis intenta alejarse de este concepto. La capacidad de intervenir en intercambios, modificar relaciones y participar en la producción de dinámicas sociales constituye el eje de la indagación. Desde este encuadre, nos concentramos en los NPC con capacidad de adaptación contextual y en otros ANHC empleados en simulaciones sociales.

Estas unidades de arquitectura multimodal que interactúan entre sí pueden generar escenarios interesantes para el estudio de sus relaciones y de las dinámicas de interacciones complejas. Cada agente tomaría decisiones basadas en su interpretación única del entorno que llevarían a fenómenos sociales como la cooperación, competencia o la formación espontánea de grupos o roles.

El objetivo del artículo es desarrollar una lectura crítico-reflexiva de los fundamentos y alcances del prototipado social con estas entidades, atendiendo a sus posibilidades heurísticas y a sus problemas epistemológicos y éticos. El trabajo adopta una metodología de análisis conceptual y revisión narrativa de bibliografía, organizada en torno a tres operaciones: reconstruir antecedentes del modelado social, examinar investigaciones que emplean agentes generativos y discutir las objeciones dirigidas a estas prácticas. Para ello, se ponen en diálogo los aportes sobre simulación de distintos autores (Gilbert & Troitzsch, 2005; Squazzoni, 2010), investigaciones (Park *et al.*, 2023, 2024; Orlando *et al.*, 2025; Yang *et al.*, 2025), y la perspectiva sobre mediación técnica y moral de Peter-Paul Verbeek (2009, 2011), articulada

con la teoría de los ANHC. Las investigaciones examinadas funcionan como casos ilustrativos para la discusión conceptual; el artículo no presenta experimentos propios ni una validación empírica de sus resultados. Su contribución específica consiste en fundamentar una interpretación del prototipado social que reconozca a los agentes como participantes en la producción de dinámicas susceptibles de investigación y permita distinguir entre estudiar las relaciones que efectivamente emergen en la simulación y extrapolar sus resultados a colectivos humanos. Esta distinción orienta la evaluación de su alcance heurístico y la discusión de sus límites.

Acerca del modelado multiagencial de interacciones sociales

En 1945 Norbert Wiener y Arthur Rosenblueth escribieron un pequeño pero importante artículo sobre la relevancia de los modelos para la ciencia (Rosenblueth & Wiener, 1945). Al menos desde entonces, no es necesario justificar su valor como herramientas heurísticas que facilitan la generación de hipótesis y la comprensión de fenómenos complejos. Tampoco es novedoso el riesgo de reificar los modelos, ni la advertencia sobre la necesidad de experimentar y testarlos en forma constante. Acaso sí puede resultar borrosa, en el medio digital, la línea entre la representación y la realidad misma².

El modelado en ciencias sociales tampoco es una metodología nueva, aunque asistimos a un cambio en las respuestas posibles a las tradicionales críticas sobre la imposibilidad de capturar la complejidad del mundo social, su adecuación epistemológica y las consecuencias normativas de su aplicación. Existen antecedentes interesantes en el uso de la computación para modelar sistemas complejos, tanto en ecología como en ciencias humanas. Cuando decimos ciencias sociales en el contexto de la teoría de los ANHC nos referimos al estudio de las redes heterogéneas en las que se entrelazan tanto actores humanos como no humanos, reconociendo que la agencia y la construcción de lo social dependen de la interacción distribuida entre estos elementos.

Uno de los pioneros en este campo fue Nils Aall Barricelli, quien trabajó con John von Neumann, utilizando tiempos ociosos de la computadora MANIAC I (Mathematical Analyzer, Numerical Integrator, and Computer), en Los Alamos, para simular procesos evolutivos en sistemas digitales. En sus experimentos, observó la formación de organismos digitales, también

² Como dice Jean-Pierre Dupuy (2009) siguiendo a von Neumann, un modelo matemático que describe un objeto complejo es en sí mismo un objeto complejo. Pasando determinado límite, los modelos no son muy diferentes de los objetos del mundo físico.

denominados simbioorganismos (Barricelli, 1962, 1963), que interactuaban, competían y cooperaban, con lo que demostró que era posible modelar procesos evolutivos en computadoras. Su trabajo fue más experimental, pero se basaba en la teoría del matemático húngaro respecto de los autómatas celulares y su posibilidad de autoreplicación (von Neumann, 1966).

En todo caso, la computarización facilita el estudio de cómo comportamientos individuales simples pueden dar lugar a fenómenos colectivos complejos. No asumimos con esto que el nivel individual tenga algún tipo de primacía ontológica, ya que, desde este enfoque, reconocemos que tanto entidades subindividuales —tales como emociones, procesos cognitivos o impulsos—, como agregados supraindividuales —tales como normas, discursos u otros vectores sociales que median y potencian la conducta— contribuyen de manera interdependiente a la emergencia de la acción colectiva.

El carácter multiescalar de los NPC evidencia los juegos entre lo individual y lo sistémico: no son solo personajes interactivos, sino un nodo dentro de un entramado mayor que abarca motores gráficos, *scripts* narrativos y dinámicas socioculturales. En su forma más rudimentaria son un conjunto de atributos individualizado que reacciona según reglas fijas; en su evolución con modelos generativos se comportan como agentes que encarnan una subjetividad sintética.

El aporte significativo de estas nuevas tecnologías, entonces, es la capacidad de integrar agentes autónomos dotados de módulos especializados. Esto no es solo una cuestión cuantitativa, pues representa un salto cualitativo en la forma en que concebimos y simulamos las dinámicas sociales. Estos agentes no humanos, al estar diseñados con habilidades específicas, como la toma de decisiones contextualizadas, la adaptación a entornos cambiantes y la interacción fluida con otros agentes (humanos o no), abren escenarios sociales complejos que antes eran inaccesibles porque resultaban demasiado simplificados o porque, en la práctica, estaban cerrados a lo impredecible.

El aporte de estas nuevas tecnologías no consiste en introducir indeterminación dentro de la simulación, sino en permitir la exploración de procesos cuya evolución resulta computacionalmente irreducible (Wolfram, 2002). Aunque las reglas locales permanezcan deterministas, la interacción entre múltiples agentes autónomos puede dar lugar a configuraciones emergentes cuya aparición no es deducible analíticamente y solo puede conocerse mediante la propia ejecución del modelo (Chaitin, 1987). La novedad, entonces, no proviene de una ruptura del determinismo sino de los límites computacionales de la predicción,

tal como ponen de manifiesto la teoría de la computabilidad inaugurada por Turing y los desarrollos posteriores sobre irreducibilidad computacional (Turing, 1937; Calude & Stay, 2008; Zenil, 2012).

En este sentido, la simulación deja de ser un simple instrumento de ilustración para convertirse en un verdadero dispositivo de descubrimiento. Si aceptamos que una simulación social rica es computacionalmente irreducible, entonces la simulación no solo representa un fenómeno: lo produce, porque el estado futuro no está contenido como conocimiento explícito en el diseño inicial, aunque sí esté determinado.

Al simular procesos cognitivos y perceptivos a nivel del agente, se obtiene una aproximación más realista de la toma de decisiones y de la interacción social, imitando la variabilidad y la impredecibilidad propias de las interacciones humanas. Estas entidades pueden resultar útiles no solo en la sociología computacional o la economía del comportamiento, sino también en la planificación urbana o incluso en la creación de entornos de entrenamiento para profesionales (pensemos en simulaciones de interacciones posibles para un trabajador social o un funcionario público).

Además, como un interés metodológico, la modularidad del sistema permite ajustar, agregar o sustituir componentes de forma independiente, lo que facilita la exploración de diferentes hipótesis y escenarios en investigaciones de ciencias sociales. Mediante este camino se puede observar y proyectar la evolución de distintas normas sociales, conflictos, dinámicas de discriminación o cooperación a gran escala.

Algunos avances ejemplares

La bibliografía clásica de las simulaciones sociales computarizadas (Lazer *et al.*, 2009; Gilbert & Troitzsch, 2005; Squazzoni, 2010) no contaba con el nivel de complejidad que podemos alcanzar en la actualidad. Si bien sus postulados siguen siendo válidos en muchos sentidos, la transformación actual no es una cuestión solo cuantitativa, de potencia computacional, sino que, ante determinado nivel de complejidad en la organización de cualquier conjunto de entidades, un sistema puede exhibir propiedades o comportamientos que no están presentes en sus componentes individuales. Es el salto que justifica cierta autonomía para cada ciencia. Por eso, ahora nos toca visitar algunos casos exitosos de aplicación de distintas técnicas de prototipado social con agentes no humanos de comunicación que cuentan con módulos de procesamiento y generación autónoma de información.

Como primer caso tenemos un estudio que modeló y simuló interacciones sociales en redes digitales entre agentes no humanos con distintas personalidades e intereses para estudiar un fenómeno contraintuitivo conocido como la paradoja de la amistad, consistente en que las personas, en promedio, tienen menos amigos en redes digitales que sus amigos (Orlando *et al.*, 2025). El fenómeno sucede por el sesgo de la media, un fenómeno estadístico condicionado por la estructura de la red. Pero lo interesante es que esta investigación logró repetir esta situación con interacciones no guiadas de ANHC, revelando mediante una simulación que existe una estructura social jerárquica donde los agentes sociales prefieren conectarse con aquellos que demuestran una mayor actividad o influencia.

En segundo lugar, presentamos un caso de estudio sobre las variables socioeconómicas que entran en juego a partir de las decisiones de ANHC que simulan actividades de agentes económicos, lo que permite entender mejor cómo las conductas individuales pueden conducir a las burbujas financieras o a las recesiones (Yang *et al.*, 2025). Esta investigación se basa en el principio de que los LLM (grandes modelos de lenguaje, por sus siglas en inglés) pueden dar cuenta de los sesgos cognitivos, las fluctuaciones emocionales y otras influencias no racionales de los agentes económicos en orden a generar simulaciones realistas que ayuden a que los decisores políticos puedan prever escenarios. El artículo publicado por este equipo cuenta cómo el desarrollo de TwinMarket, un novedoso marco multiagencial desarrollado por ellos mismos, permite prever cómo determinados fenómenos de comportamiento colectivo se ven impulsados según ciertas condiciones de conductas individuales. Incluye variables como las creencias, deseos e intenciones de los agentes no humanos vinculados a microacciones socioeconómicas, tales como leer anuncios y noticias, tendencias en redes sociales y recomendaciones de acciones de distintos actores de los mercados existentes.

El tercer caso es uno de los más difundidos, un equipo de investigadores, principalmente de la Universidad de Stanford, desarrolló una pequeña villa de 25 ANHC inspirada en el popular videojuego The Sims (Park *et al.*, 2023). Su arquitectura ofrece una emulación verosímil de intercambios sociales humanos y diseñaron un conjunto de medidas descriptivas para analizar los comportamientos sociales emergentes. Estas medidas se centran en tres fenómenos particulares: la difusión de información, la formación de relaciones y la coordinación de los agentes. Por ejemplo, programaron en la memoria de un agente la organización de una fiesta por el día de San Valentín y en otro el lanzamiento de una campaña política personal para ser una autoridad local. Dos días después, midieron el grado de alcance de ambas propuestas en el resto de los agentes, quiénes se habían enterado y quiénes no, de

estas novedades. Incluso más, respecto de la fiesta, hubo otros agentes que organizaron un encuentro para preparar la celebración, y está el registro de quienes asistieron. Por lo tanto, tanto la difusión de la información, la construcción de redes entre los agentes y la colaboración entre ellos pudieron ser objetos de medición.

Un año después elevaron a mil el número de agentes participantes de su base de datos, útiles para un proyecto similar (Park *et al.*, 2024), pero esta vez reclutaron para ello a más de mil seres humanos estadounidenses y les realizaron entrevistas en profundidad, de más de siete mil palabras, que sirvieran como insumo para el modelado de los ANHC con LLM. Dos semanas después de la primera entrevista midieron la predictibilidad de cada ANHC para anticipar las respuestas de su modelo humano a encuestas abiertas, mediante técnicas tradicionales de análisis sociológico, y encontraron un resultado de más del 85 % de exactitud en estas. Es decir, es altamente predecible la actitud y el comportamiento de los humanos, al menos a corto plazo, argumento útil para apoyar la eficacia del prototipado social en contextos restringidos.

El ombligo de mi mejor amigo

Los NPC, el caso más común de estos agentes, funcionan como nodos comunicacionales dentro de un ecosistema híbrido, donde coexisten con usuarios humanos y otros ANHC (LLMs, bots, asistentes virtuales, etc.). Purnendu Mukherjee, CEO de Convai, una plataforma para crear personajes 3D útiles en el desarrollo de videojuegos y mundos virtuales, está seguro de que estos pueden convertirse en los mejores amigos de los usuarios (Sidhu, 2024). Mukherjee menciona su propia experiencia de socialización con agentes no humanos sin módulos generativos, apenas con algunas entradas configuradas, que funcionaron como una cortina amable ante un entorno humano hostil, por su cualidad de extranjero en Estados Unidos.

Desde nuestro enfoque posfenomenológico, las tecnologías no son entendidas como meros instrumentos, sino como mediaciones que coconfiguran la experiencia humana. Los NPC no solo responden a los jugadores, enmarcan el mundo y condicionan cómo se experimenta tanto la propia inmersión como el resto de las relaciones del usuario. Las experiencias con agentes no humanos, que brindan tantas empresas hoy, significan una comunicación realista y adaptada a las necesidades de los usuarios, que les proporciona compañía y apoyo emocional constante. Que los agentes interactivos sean autómatas replicados no les impide fungir de acompañantes.

Ambos tipos de agentes se mezclan y puede resultar difícil distinguir la conducta de unos y otros, en especial cuando los investigadores solo cuentan con el trazo digital de ambos. Un grupo de estudio del Carnegie Mellon estima que los bots son responsables de, al menos, el 20 % del tráfico digital (Ng & Carley, 2025).

Existen modelos que combinan técnicas de análisis de emociones, NLP y redes neuronales para detectar bots en redes sociales (Castillo Cárdenas & Ávila Camacho, 2024) o en el sistema financiero (Bendada *et al.*, 2025), pero no están demasiado cerca de lograr una eficacia absoluta en su cometido. Es decir, ni los humanos ni los ANHC podemos garantizar una distinción absoluta en la identificación del origen de la agencia. Hablamos de, en su mayoría, agentes a detectar que no cuentan con módulos especializados. Otro equipo de investigación chino (Qiao *et al.*, 2024) desarrolló uno de estos agentes con un módulo de LLM que pudo sortear muy bien los métodos tradicionales de detección de bots e involucrarse en interacciones sociales modelando escenarios de flujos de información e intercambios con otros usuarios. Por lo cual es aún más difícil registrar una diferencia en estos casos.

Sin lugar a dudas, la dimensión más relevante para entender la creciente imbricación entre agentes híbridos es la antropomorfización. Que los ANHC simulen o parezcan ser humanos constituye lo que Bernard Stiegler llamaría *pharmakon*, remedio y veneno a la vez (Stiegler, 2002). En otro trabajo desarrollé esta idea, sopesando los peligros derivados de la confusión antropomórfica (Tagnin, 2024b), pero en este quiero detenerme en algunas causas de su eficiencia. ¿Por qué los usuarios pueden desarrollar vínculos de confianza, intimidad o compañía con agentes no humanos, aun sabiendo que no son humanos?

Para los investigadores mexicanos Duque Parra y Santes Ortega la antropomorfización constituye el nivel más alto de interacción emocional posible, y ya hay autores que clasificaron distintas formas de manifestación de este fenómeno (Ibrahim *et al.*, 2025; Ma *et al.*, 2025). Los resultados de su investigación sugieren que el compromiso emocional se desarrolla con el tiempo, el uso y la profundidad creciente del intercambio informativo (Duque Parra & Santes Ortega, 2026).

La disposición humana a establecer vínculos con agentes conversacionales precede a los modelos generativos contemporáneos. Ya Weizenbaum observó que muchas personas atribuían comprensión y empatía a uno de los primeros intentos de superar el test de Turing, un programa muy simple (Weizenbaum, 1966).

La creciente confianza depositada en NPC y otros ANHC no puede explicarse solo como un error cognitivo o una incapacidad para distinguir humanos de máquinas. Las personas

tienden a aplicar en forma espontánea reglas sociales a los sistemas interactivos, son educados en su trato con ellos, les asignan personalidad y reaccionan emocionalmente a ellos (Reeves & Nass, 1996).

Sherry Turkle enseña que muchas personas buscan características como la escucha, la disponibilidad y la sensación de compañía en ANHC debido a que se dan en un marco de ausencia de conflictos y sensación de control sobre la interacción (Turkle, 2011, 2015). Sostiene que los robots y agentes conversacionales, al igual que las mascotas, prometen una relación donde uno puede sentirse acompañado sin experimentar la complejidad de la reciprocidad humana. Tal como atestiguaba Mukherjee, el NPC funciona como refugio frente a experiencias sociales percibidas como hostiles. Esto se vincula con la interpretación de esta relación por los aportes de la teoría del apego. Para la misma, los usuarios buscan seguridad emocional, validación, presencia constante y espacios de autoexpresión, muchas veces para combatir la soledad (Cassidy & Shaver, 2016).

En todo caso, la construcción de vínculos con agentes no humanos tiene fortalezas que auguran una gran proyección. Ya muchos han cruzado con éxito el *uncanny valley*³ (Mori, 2012) con los ANHC.

Reacciones anticipadas a críticas humanistas

Diferentes corrientes dentro de las ciencias sociales (en especial aquellas influenciadas por la teoría crítica, los estudios culturales y otras perspectivas humanistas) han planteado críticas al uso de modelos para representar las interacciones sociales. Sus principales vectores apuntan al reduccionismo y a la simplificación que esto supondría, junto a la concomitante omisión de contextos culturales o históricos únicos.

Otra crítica es aquella dirigida hacia el riesgo de perpetuar o incluso legitimar determinadas estructuras de poder que supone ignorar las dinámicas de dominación y resistencia que subyacen a la formalización de las interacciones mediante algoritmos o reglas predefinidas. Existe además una disputa sobre los marcos interpretativos que dan sentido a las

³ El concepto de *uncanny valley*, traducido como “valle inquietante”, fue propuesto por el ingeniero japonés Masahiro Mori (2012) para describir el fenómeno por el cual un aumento en la semejanza antropomórfica de un robot con los seres humanos puede generar al inicio una respuesta positiva, pero, al alcanzar un grado de similitud muy elevado sin llegar a una apariencia plenamente humana, produce una sensación de extrañeza, incomodidad o rechazo.

instituciones y a las conductas de los agentes. Llamémosle la objeción sobre la dimensión político-epistémica de los modelos constructivos.

Las últimas críticas que vamos a tener presente son aquellas que podemos agrupar en la categoría que considera irremplazable la ambigüedad moral de las interacciones humanas. Una acusada falta de reflexividad y autocrítica haría imposible modelar la esencia de las interacciones sociales humanas, su dimensión ética y excepcional. Estos tres grandes grupos de críticas no buscan ser exhaustivos ni mutuamente excluyentes. Cualquier argumentación en contra de la simetría agencial propuesta, en general, por la teoría de los Agentes no Humanos de Comunicación para el conjunto de objetos de estudio definido por esta (Tagnin, 2024a) suele hacer uso de razones que atraviesan las categorías mencionadas. Para intentar ordenar el estado de situación, vamos a anticiparnos con algunos reparos.

Ante las objeciones de reduccionismo y simplificación, cuyo principal argumento sería algo así como que estos modelos reducen la complejidad humana a variables cuantificables que obliteran la riqueza de las experiencias subjetivas, podemos decir que en la búsqueda de representar interacciones sociales, los modelos pueden ser diseñados de manera flexible, y no como una representación estricta de una realidad única. Los modelos pueden evolucionar y adaptarse a contextos específicos, y se pueden incorporar aspectos culturales e históricos mediante la personalización de los agentes, como en el experimento de Park. La necesaria simplificación de variables que implica un modelo no significa un desentendimiento total de los contextos. De hecho, se amplifica la complejidad al explorar distintos potenciales. El propósito de la simplificación es ofrecer un punto de partida para estudios más profundos y tener un espacio de pruebas dentro de un marco controlado, pero el tejido de actitudes y comportamientos puede resultar mucho más complejo incluso que en un escenario “natural”.

Sobre el riesgo de ser funcional a ciertas estructuras de poder, es necesario advertir que los algoritmos y las reglas predefinidas no son incapaces de abordar estas cuestiones, máxime cuando la complejidad de los módulos especializados tiene la posibilidad de enseñar escenarios que no imaginó ningún humano previamente. Muchos modelos pueden simular situaciones de conflicto, desigualdad y resistencia para observar cómo las interacciones evolucionan bajo diversas condiciones de relaciones de poder. Pero de manera aún más interesante, los mismos ANHC pueden contar con la capacidad de reflexionar sobre su entorno y actuar en consecuencia, generando dinámicas originales de resistencia, lo que podría incluso ayudar a ilustrar cómo las acciones sociales pueden reproducir o transformar esas estructuras de poder.

Como corolario, podrían crear un espacio para cuestionar el mismo proceso de modelado social.

Debemos ahora analizar lo que hemos llamado la objeción sobre la dimensión político-epistémica de los modelos constructivos. La elección de las variables y características de las entidades nunca es neutral. Cada modelado encarna supuestos sobre el mundo social y produce efectos políticos y epistémicos. Sin embargo, nadie espera que los modelos sean capaces de eliminar la tensión entre marcos interpretativos, pero podemos señalar su utilidad para la formalización de estos y la consecuente evaluación sobre cómo distintas asunciones impactan en los resultados. Por ejemplo, podemos pensar cómo un modelo funcionalista y uno crítico pueden coexistir, y proporcionar distintas formas de interpretar y predecir la realidad.

Por último, tenemos que atender la advertencia sobre la supuesta irremplazabilidad humana en cuestiones morales. Lo primero que vamos a considerar es que el tipo de prototipado por el que abogamos representa elementos que permiten reflexionar sobre los dilemas éticos a través de las interacciones que emergen. La carga moral de los objetos técnicos está sobradamente discutida en el ámbito de la filosofía de la técnica (Heidegger, 2021; Himma, 2009; Floridi, 2014; Ortega y Gasset, 1964; Anders, 2011, entre otros). De hecho, la mediación tecnológica de la moralidad es un fenómeno bien documentado por la filosofía neerlandesa. Desde las condiciones en las que la materialidad de los artefactos condiciona la acción (Verbeek, 2009, 2011) hasta la posibilidad de atribuir responsabilidad distribuida en sistemas sociotécnicos (van de Poel, 2023), el interés de esta rama disciplinar no se limita a la deliberación individual, sino que se articula a partir del análisis de las infraestructuras que orientan el tipo de decisiones morales que nos interesan. En este contexto, los modelos con ANHC pueden servir no solo para mapear estas mediaciones, sino también para explorar sus posibles reconfiguraciones.

Esto permitiría analizar, por ejemplo, cómo emergen patrones de responsabilidad en sistemas híbridos de humanos y máquinas, y hacer explícitas las configuraciones normativas que ya operan en la interacción entre humanos y tecnologías. Al contrario de lo que dicen quienes piensan que este método podría encubrir responsabilidades humanas, o que puede ocultar intereses, creemos que el prototipado social puede ayudarnos a tomar conciencia de variables incidentes relevantes y, muchas veces, ignoradas.

Así, en lugar de ver la ambigüedad de las consideraciones éticas de los agentes como algo que no puede ser modelado, se pueden entender como un componente activo y presente dentro de los escenarios planteados, que podrían significar una instancia válida para

profundizar la comprensión de los valores que asignamos los humanos. Esto nos lleva a la última cuestión: el romanticismo desde el cual se argumenta que existe una excepcionalidad humana irremplazable, cierta esencialidad de las interacciones sociales, ya sea por la interpretación de significados, la construcción de identidades o la negociación de valores, que solo nosotros, humanos, poseemos. Sin embargo, esta supuesta excepcionalidad humana no es un dato empírico, sino un postulado filosófico cuyos matices han sido refutados en múltiples ocasiones a medida que avanzan las ciencias cognitivas, la inteligencia artificial y la filosofía de la tecnología (Dehaene & Naccache, 2001; Dehaene *et al.*, 2021; Kohn, 2013; Foucault, 1966; Gunkel, 2012). Si definimos la interacción social en términos de intercambio de información, construcción de sentido y regulación de comportamiento, los ANHC, y los NPC en particular, ya participan en estos procesos.

Para comprender mejor estos fenómenos es necesario evitar el antropocentrismo de las concepciones que sostienen nuestra excepcionalidad en esos términos. Existen varios esfuerzos conceptuales que van en ese sentido (Tagnin, 2024a, 2026a, 2026b).

Conclusiones

El avance de los agentes no humanos en entornos digitales plantea una reconfiguración profunda de los marcos epistemológicos desde los cuales las ciencias sociales comprenden la acción, la agencia y la subjetividad. A través del estudio de los NPC como nodos relacionales activos dentro de ecologías sociotécnicas, se hace evidente la necesidad de superar las dicotomías clásicas entre sujeto y objeto, humano y máquina, natural y artificial.

El prototipado social se revela en este contexto como una herramienta metodológica y como una práctica heurística que permite experimentar con configuraciones posibles de lo social. Esta práctica permite observar cómo se estabilizan (o no) formas emergentes de normatividad, reconocimiento, alteridad y agencia, en sistemas híbridos donde la acción es distribuida entre humanos y artefactos dotados de cierto grado de autonomía simbólica o performativa.

Los modelos de NPC o de otros ANH no son representaciones estáticas y definitivas de la realidad social, sino que actúan en una constante evolución y permiten probar, cuestionar y desafiar diversas formas de interacción social. Lejos de ser una forma de reduccionismo, estos modelos abren espacios para explorar dinámicas complejas de interacción que no podrían ser observadas de otra manera.

En este sentido, el prototipado social basado en ANHC no debe entenderse solo como un nuevo recurso metodológico, sino como la expresión de una transformación más amplia en la producción del conocimiento social. Así como las ciencias naturales incorporaron instrumentos que ampliaron las capacidades perceptivas del investigador, las ciencias sociales comienzan hoy a disponer de dispositivos capaces de explorar dinámicas interactivas cuya complejidad excede la intuición humana. Estos modelos amplían el horizonte de hipótesis posibles, permiten experimentar con configuraciones contrafácticas y ofrecen un espacio controlado para observar la emergencia de fenómenos colectivos. No anulan la tradición interpretativa de las ciencias humanas, sino que la potencian.

Los ANHC transforman la propia epistemología de las ciencias sociales porque modifican qué puede conocerse y cómo puede conocerse. Estamos frente a una ampliación de las capacidades heurísticas de las ciencias sociales, coherente con la teoría de los ANHC y con una concepción posfenomenológica de la técnica, frente a un cambio de paradigma metodológico que permite estudiar nuevos objetos epistémicos. Las simulaciones dejan de ser representaciones estáticas para convertirse en espacios donde pueden emerger fenómenos sociales que no estaban previstos por el investigador.

El desafío ya no consiste en decidir si estos agentes pueden participar de la investigación social, sino en desarrollar los marcos epistemológicos, metodológicos y éticos que permitan aprovechar ese potencial. La agenda futura debería desarrollarse en torno a los problemas que ahora quedan planteados: ¿cómo validar este tipo de prototipos?, ¿qué criterios permiten evaluar su poder explicativo?, ¿cómo evitar que reproduzcan sesgos del diseño?, ¿qué límites éticos aparecen cuando los agentes adquieren mayor autonomía?, ¿cómo integrar estos modelos con metodologías cualitativas tradicionales? Son algunas de las preguntas que quedan por resolver.

Referencias

- Anders, G. (2011). *La obsolescencia del hombre*. Pre-Textos.
- Barricelli, N. A. (1962). Numerical testing of evolution theories: Part I. Theoretical introduction and basic tests. *Acta Biotheoretica*, 16(1–2), 69–98
<https://doi.org/10.1007/BF01556771>
- Barricelli, N. A. (1963). Numerical testing of evolution theories: Part II preliminary tests of performance. Symbiogenesis and terrestrial life. *Acta Biotheoretica*, 16(3–4), 99–126.
<https://doi.org/10.1007/BF01556602>
- Bendada, A. M. R., Korba, A. A., Bouchiha, M. A., & Ghamri-Doudane, Y. (2025). *BotDetect: A decentralized federated learning framework for detecting financial bots on the EVM blockchains (Version 1) [Preprint]*. arXiv.
<https://doi.org/10.48550/ARXIV.2501.12112>
- Calude, C. S., & Stay, M. A. (2008). Most programs stop quickly or never halt. *Advances in Applied Mathematics*, 40(3), 295–308. <https://doi.org/10.1016/j.aam.2007.01.001>
- Cassidy, J., & Shaver, P. R. (Eds.). (2016). *Handbook of attachment: Theory, research, and clinical applications* (3rd ed.). Guilford Press.
- Castillo Cárdenas, S. P., & Ávila Camacho, F. J. (2024). Redes Neuronales Recurrentes para la detección de noticias falsas y cuentas bot en Twitter. *RICT Revista de Investigación Científica, Tecnológica e Innovación*, 2(4), 64–71.
<https://doi.org/10.5281/zenodo.14194198>
- Chaitin, G. J. (1987). *Algorithmic information theory*. Cambridge University Press.
- Dehaene, S., Lau, H., & Kouider, S. (2021). What is consciousness, and could machines have it? En J. von Braun, M. S. Archer, G. M. Reichberg, & M. Sánchez Sorondo (Eds.), *Robotics, AI, and humanity* (pp. 43–56). Springer. https://doi.org/10.1007/978-3-030-54173-6_4
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1–2), 1–37.
[https://doi.org/10.1016/S0010-0277\(00\)00123-2](https://doi.org/10.1016/S0010-0277(00)00123-2)
- Dupuy, J. P. (2009). *On the origins of cognitive science: The mechanization of the mind*. MIT Press.
- Duque Parra, J. P., & Santes Ortega, A. (2026). Symmetries and asymmetries between attitudes and interaction in relation to the emotional uses of LLMs. *AI & Soc*, 1–19.
<https://doi.org/10.1007/s00146-026-03069-8>

- Floridi, L. (2014). Artificial agents and their moral nature. En P. Kroes & P.-P. Verbeek (Eds.), *The moral status of technical artefacts* (pp. 185–212). Springer.
- Foucault, M. (1966). *Las palabras y las cosas: Una arqueología de las ciencias humanas*. Siglo XXI.
- Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist*. Open University Press.
- Gunkel, D. (2012). *The machine question. Critical perspectives on AI, robots, and ethics*. MIT Press. <https://doi.org/10.7551/mitpress/8975.001.0001>
- Heidegger, M. (2021). *La pregunta por la técnica*. Herder Editorial.
- Himma, K. (2009). Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent? *Ethics and Information Technology*, 11(1), 19–29. <https://doi.org/10.1007/s10676-008-9167-5>
- Ibrahim, L., Akbulut, C., Elasmara, R., Rastogi, C., Kahng, M., Morris, M. R., McKee, K. R., Rieser, V., Shanahan, M., & Weidinger, L. (2025). *Multi-turn evaluation of anthropomorphic behaviours in large language models*. arXiv. <https://doi.org/10.48550/arXiv.2502.07077>
- Kohn, E. (2013). *How forests think: Toward an anthropology beyond the human*. University of California Press. <https://doi.org/10.1525/california/9780520276109.001.0001>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science*, 323(5915), 721–723. <https://doi.org/10.1126/science.1167742>
- Ma, N., Khynevych, R., Hao, Y., & Wang, Y. (2025). Effect of anthropomorphism and perceived intelligence in chatbot avatars of visual design on user experience: Accounting for perceived empathy and trust. *Frontiers in Computer Science*, 7, Article 1531976. <https://doi.org/10.3389/fcomp.2025.1531976>
- Ng, M. (2012). The uncanny valley [from the field] (K. F. MacDorman & N. Kageki, Trans.). *IEEE Robotics & Automation Magazine*, 19(2), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Ng, L. H. X., & Carley, K. M. (2025). A global comparison of social media bot and human characteristics. *Scientific Reports*, 15(1). Article 10973. <https://doi.org/10.1038/s41598-025-96372-1>

- Orlando, G. M., La Gatta, V., Russo, D., & Moscato, V. (2025). Can generative agent-based modeling replicate the friendship paradox in social media simulations. En *Proceedings of the 17th ACM Web Science Conference (WebSci '25)* (pp. 510–515). Association for Computing Machinery. <https://doi.org/10.1145/3717867.3717895>
- Ortega y Gasset, J. (1964). *Meditación de la técnica*. Revista de Occidente.
- Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. En *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (Article 2, pp. 1–22). Association for Computing Machinery. <https://doi.org/10.1145/3586183.3606763>
- Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C., Morris, M. R., Willer, R., Liang, P., & Bernstein, M. S. (2024). *Generative agent simulations of 1,000 people*. arXiv preprint arXiv:2411.10109. <https://scispace.com/papers/generative-agent-simulations-of-1000-people-31sj9byjex26>
- Qiao, B., Li, K., Zhou, W., Li, S., Lu, Q., & Hu, S. (2024). *BotSim: LLM-powered malicious social botnet simulation*. arXiv. <https://doi.org/10.48550/arXiv.2412.13420>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.
- Rosenblueth, A., & Wiener, N. (1945). The role of models in science. *Philosophy of Science*, 12(4), 316-321. <https://doi.org/10.1086/286874>
- Sidhu, B. (Host). (2024, junio 18). Could your new best friend be an AI-powered NPC? w/ Purnendu Mukherjee [episodio de podcast]. En *The TED AI Show*. TED. https://www.ted.com/talks/the_ted_ai_show_could_your_new_best_friend_be_an_ai_powered_npc_w_purnendu_mukherjee
- Squazzoni, F. (2010). The impact of agent-based models in the social sciences after 15 years of incursion. *History of economic ideas*, 18(2), 197-233 <https://doi.org/10.1400/148280>
- Stiegler, B. (2002). *La técnica y el tiempo: Vol. 1. El pecado de Epimeteo*. Hiru.
- Tagnin, J. (2024a). Sobre la relevancia moral de los agentes no humanos de comunicación (ANHC). *Ti. Futuros comunes-Revista de Tecnologías informacionales* (4), 19-28. <https://publicaciones.unpaz.edu.ar/OJS/index.php/ti/article/view/1752>
- Tagnin, J. (2024b). No tan humanos. Acerca de la antropomorfización de los agentes no humanos de comunicación. *Bordes*, (34), 37-46. <https://revistabordes.unpaz.edu.ar/no->

[tan-humanos-acerca-de-la-antropomorfizacion-de-los-agentes-no-humanos-de-comunicacion/](#)

- Tagnin, D. J. (2026a). Consideraciones éticas sobre los agentes no humanos de comunicación como productores científicos. *Revista de Investigación del Departamento de Humanidades y Ciencias Sociales*, (29), 123-145. <https://doi.org/10.54789/rihumso.26.15.29.7>
- Tagnin, D. J. (2026b). Alienación y teoría de los agentes no humanos de comunicación: un diálogo posible con las propuestas de la entrañabilidad tecnológica. *Ucronías*, (13), 13-28. <https://doi.org/10.5281/zenodo.20819153>
- Turing, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2), 230–265. <https://doi.org/10.1112/plms/s2-42.1.230>
- Turkle, S. (2011). *Alone together. Why we expect more from technology and less from each other*. Basic Books.
- Turkle, S. (2015). *Reclaiming conversation: The power of talk in a digital age*. Penguin.
- van de Poel, I. (2023). AI, control and unintended consequences: The Need for Meta-Value. In *Philosophy of Engineering and Technology* (pp. 117–129). Springer International Publishing. https://doi.org/10.1007/978-3-031-25233-4_9
- van de Poel, I. (2023). AI, control and unintended consequences: The need for meta-values. En A. Fritzsche & A. Santa-María (Eds) *Rethinking Technology and Engineering. Philosophy of Engineering and Technology*, vol 45 (pp. 117–129). Springer. https://doi.org/10.1007/978-3-031-25233-4_9
- Verbeek, P. P. (2009). Moralizing technology: On the morality of technological artifacts and their design. En D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 226–243). Rowman & Littlefield.
- Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.
- von Neumann, J. (1966). *Theory of self-reproducing automata* (A. W. Burks, Ed.). University of Illinois Press.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Wolfram, S. (2002). *A new kind of science*. Wolfram Media.

- Yang, Y., Zhang, Y., Wu, M., Zhang, K., Zhang, Y., Yu, H., Hu, Y., & Wang, B. (2025). *TwinMarket: A scalable behavioral and social simulation for financial markets* (Version 5) [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2502.01506>
- Zenil, H. (Ed.). (2012). *A computable universe: Understanding and exploring nature as computation*. World Scientific.