

La Frontera de la Discriminación Algorítmica: Análisis Exhaustivo del Caso iTutorGroup y el Futuro de la Responsabilidad Civil en la IA

1. Introducción: El Cambio de Paradigma en el Reclutamiento Automatizado

Desde la perspectiva analítica que define la "Firma Scarpa" en derechoartificial.com, hemos insistido en que la Inteligencia Artificial (IA) no es una herramienta aséptica, sino una *machina económica* que actúa como un prisma de las estructuras sociales preexistentes. El caso *EEOC v. iTutorGroup* no debe ser leído como un incidente aislado, sino como el hito fundacional de la **frontera de la responsabilidad algorítmica**. Nos encontramos ante el nacimiento de una nueva era de litigiosidad laboral donde la opacidad técnica ya no sirve de refugio contra el escrutinio de los derechos civiles.

Siguiendo la investigación de **Sheard et al. (2025)**, debemos conceptualizar este fenómeno como "**discriminación algorítmica facilitada**". No se trata de un error técnico accidental, sino de un problema sistémico que emerge de un ensamblaje complejo de actores humanos y no humanos. En este ecosistema, la infraestructura tecnológica reduce drásticamente los costos de discriminar, otorgando un halo de objetividad matemática a decisiones sesgadas. La iniciativa de la EEOC (*Artificial Intelligence and Algorithmic Fairness Initiative*) confirma que el *onus probandi* y la vigilancia del Título VII y la ADEA se extienden con vigor sobre estos sistemas, obligándolos a diseccionar los hechos probados para blindar la equidad en el mercado de trabajo.

2. Anatomía del Caso EEOC v. iTutorGroup: Hechos, Descubrimiento y Sanción

Este litigio constituye un punto de inflexión estratégico al representar el primer acuerdo (*settlement*) de la EEOC en materia de discriminación por IA. La relevancia del caso trasciende las fronteras estadounidenses, pues subraya que la jurisdicción laboral alcanza a la nube digital independientemente del *locus* del empleador.

Elementos Diferenciadores y Hechos Probados

- **La mecánica del sesgo deliberado:** A diferencia de otros sistemas donde el sesgo es una consecuencia imprevista, iTutorGroup programó explícitamente su software de reclutamiento para ejecutar un rechazo automático basado en criterios de edad. Se excluyó sistemáticamente a mujeres mayores de 55 años y a hombres mayores de 60 años, afectando a más de 200 profesionales cualificados.

- **Alcance Multinacional y Jurisdicción:** El caso involucró a entidades radicadas en el extranjero —**Shanghai Ping'An Intelligent Education Technology Co. Ltd.** y **Tutor Group Limited (Hong Kong)**— que contrataban tutores remotos basados en EE. UU. La lección estratégica es clara: las fronteras internacionales no operan como un escudo frente a la ADEA o el Título VII cuando la prestación del servicio ocurre en suelo estadounidense.
- **El método de detección (Experimento de Doble Solicitud):** La discriminación fue expuesta mediante una técnica de *testing* jurídico: una aspirante (Wendy Pincus) fue rechazada tras ingresar su fecha de nacimiento real; no obstante, al presentar una solicitud idéntica al día siguiente con una fecha más reciente, el sistema le ofreció una entrevista de inmediato.
- **Remediación y Consecuencias:** El acuerdo impuso una sanción de **\$365,000** en concepto de salarios atrasados y daños. Más allá del impacto pecuniario, la empresa se sometió a una arquitectura de cumplimiento estricta: auditorías de la EEOC durante cinco años, adopción de políticas antidiscriminatorias y la obligación de invitar a los candidatos rechazados a postularse nuevamente.

Evaluación Crítica: El uso de software de terceros no constituye una "eximente" de responsabilidad. En la cadena de suministro de RR.HH., el empleador retiene el control y, por ende, la responsabilidad final. La automatización no es una patente de corso para eludir la ley.

3. Dicotomía Jurídica: Trato Dispar (Intentional) vs. Impacto Dispar (Facialmente Neutral)

En la jurisprudencia estadounidense, la discriminación algorítmica se bifurca en dos teorías cardinales que definen la estrategia del litigante.

Análisis Comparativo de Teorías de Responsabilidad

Categoría	Caso Referencia de	Mecanismo de Discriminación
Trato Dispar (Intencional)	<i>EEOC v. iTutorGroup</i>	Programación explícita de criterios discriminatorios. La intención humana se codifica directamente en el algoritmo.
Impacto Dispar (Facialmente Neutral)	<i>Mobley v. Workday, Inc.</i>	Algoritmos neutrales que generan resultados sesgados mediante variables <i>proxy</i> (ej. códigos postales correlacionados con raza o patrones de voz vinculados a discapacidades).

El desafío reside en la "**Black Box**" o caja negra. Mientras que en el trato dispar hay una instrucción clara, en el impacto dispar la discriminación es a menudo **endógena**: el sistema aprende de datos históricos sesgados y replica patrones de exclusión de forma automatizada. Casos como *Mobley v. Workday* demuestran que incluso sin mencionar la raza o la edad, el uso

de *proxies* (como el código postal para la raza o el historial educativo para la edad) crea bucles de retroalimentación que dificultan enormemente la carga de la prueba para el demandante ante la opacidad computacional.

4. Propuestas de Reforma Legislativa: La Enmienda LeRoy al Título VII

Ante la obsolescencia de las normativas vigentes, el profesor **Michael H. LeRoy** propone una solución pragmática: no intentar regular la volatilidad del código, sino redefinir las categorías legales de responsabilidad.

Su propuesta central consiste en añadir **diecisésis palabras estratégicas** a la Sección 701(c) de la Ley de Derechos Civiles (Título VII), redefiniendo a los proveedores de IA como "agencias de empleo" (*employment agencies*). La enmienda integraría el siguiente lenguaje:

"...and includes an agent of such a person, including any person who uses 'artificial intelligence' or 'AI' within the meaning set forth in 15 U.S.C. 9401(3)."

Evaluación Crítica: Esta redefinición es revolucionaria. Al transformar a las tecnológicas de Silicon Valley en "agencias de empleo" ante la ley, se les despoja de su estatus de meros proveedores de software. Esto permite aplicar décadas de jurisprudencia sobre la responsabilidad de terceros y establece un régimen de **responsabilidad directa** (o incluso objetiva), impidiendo que los desarrolladores eludan su deber de diligencia sobre los sesgos que inyectan en el mercado laboral.

5. Perspectiva Económica: Ineficiencias del Mercado y el Modelo de Carolina Pinasco

Desde el Análisis Económico del Derecho, la discriminación algorítmica es un fallo de mercado sistemático. El modelo de **Carolina Pinasco** sostiene que la mitigación de sesgos no es solo un imperativo ético, sino una condición de eficiencia económica.

1. **Asimetría de Información de Segundo Orden:** Aplicando el "mercado de limones" de Akerlof, el algoritmo se convierte en el "limón". La opacidad de la IA impide que los actores evalúen la calidad de las decisiones. Si el sistema es un "limón" sesgado, el mercado laboral colapsa al rechazar talento productivo basándose en criterios erróneos.
2. **Externalidades Negativas:** Las empresas que despliegan IA sesgada no internalizan los costos de sus errores. El costo del "falso negativo" (el trabajador cualificado rechazado injustamente) es trasladado al sistema judicial y a la sociedad en forma de ineficiencia y desempleo.
3. **El Principio del "Evitador de Menor Costo" (Least-Cost Avoider):** Económicamente, la responsabilidad debe recaer en quien puede prevenir el daño al menor costo. El desarrollador, durante la fase de diseño y entrenamiento, es quien tiene la capacidad

técnica de auditar el código. Obligar a los proveedores a internalizar este riesgo incentiva una innovación de mayor calidad.

6. Gobernanza y Futuro: Hacia una Arquitectura de Auditoría Dual

El futuro de la gobernanza algorítmica debe integrar las lecciones de la **NYC Local Law 144** y marcos como la **AIVIA** de Illinois. La gestión estratégica del riesgo exige trascender la auditoría técnica superficial para abrazar una validación empírica y económica.

1. **Auditorías de Sesgo por Terceros:** Evaluaciones independientes obligatorias que cumplan con estándares de transparencia pública (modelo NYC Local Law 144).
2. **Análisis de Impacto Dispar Interseccional:** Evaluación de cómo se cruzan variables protegidas (ej. raza y género simultáneamente), siguiendo el estándar de *Gender Shades*.
3. **Validación de Constructo:** Demostración empírica de que los criterios del algoritmo (ej. patrones de voz o micro-expresiones) miden realmente elementos esenciales del desempeño laboral.
4. **Implementación de Model Cards y Trazabilidad:** Documentación estandarizada que detalle el origen de los datos de entrenamiento y las limitaciones del modelo.
5. **Cuantificación de Falsos Negativos:** Análisis de impacto económico para medir el talento perdido por errores del sistema.
6. **Transparencia Explicable (XAI):** Traducir los resultados de la "Caja Negra" en criterios inteligibles para candidatos y reguladores.
7. **Privilegio Abogado-Cliente:** Realizar las auditorías internas bajo supervisión legal para proteger los hallazgos sensibles de futuras acciones de descubrimiento.
8. **Eliminación de Proxies Endógenos:** Auditoría constante de variables neutrales que actúan como sustitutos de categorías protegidas.
9. **Gobernanza Proactiva de Datos:** Asegurar que los datos históricos no codifiquen prejuicios del pasado antes de ser injectados al modelo.
10. **Notificación y Consentimiento Informado:** Cumplimiento estricto de leyes como AIVIA, informando a los candidatos sobre el uso de análisis algorítmico en entrevistas de video.

Como juristas en la era de la *machina económica*, nuestra responsabilidad es asegurar que la eficiencia tecnológica no se convierta en una máscara para la injusticia estructural. La ley debe evolucionar con la velocidad del algoritmo para garantizar que la innovación sea, ante todo, una herramienta al servicio de la equidad y la sostenibilidad económica.

Ricardo Scarpa Director de derechoartificial.com