

# Herramientas de predicción de violencia basada en género y feminicidio mediante la Inteligencia Artificial<sup>1</sup>

*Tools for predicting gender-based violence and femicide using Artificial Intelligence*

Marcela del Pilar Roa Avella<sup>2</sup> , Jesús Eduardo Sanabria-Moyano<sup>3</sup>  & Katherin Dinashurtado<sup>4</sup> 

Universidad Militar Nueva Granada - Colombia



**Para citaciones:** Roa Avella, M., Sanabria Moyano, J., & Dinashurtado, K. (2023). Herramientas de predicción de violencia basada en género y feminicidio mediante la Inteligencia Artificial. *Revista Jurídica Mario Alario D'Filippo*, 15(30), 360-390. <https://doi.org/10.32997/10.32997/2256-2796-vol.15-num.30-2023-4254>

**Recibido:** 22 de octubre de 2022

**Aprobado:** 30 de enero de 2023

**Editor:** Jorge Payares Bossa. Universidad de Cartagena-Colombia.

**Copyright:** © 2023. Roa Avella, M., Sanabria Moyano, J., & Dinashurtado, K. Este es un artículo de acceso abierto, distribuido bajo los términos de la licencia <https://creativecommons.org/licenses/by-nc-nd/4.0/> la cual permite el uso sin restricciones, distribución y reproducción en cualquier medio, siempre y cuando que el original, el autor y la fuente sean acreditados.



## RESUMEN

La Violencia basada en género, ha sido definida por la Organización de las Naciones Unidas como cualquier acto dañino basado en las diferencias de género atribuidas socialmente. Dentro de sus muchas manifestaciones, aquella violencia en su máxima expresión llega hasta el feminicidio; fenómeno que lejos de disminuir, se ha extendido alrededor del mundo.

Por otra parte, la inteligencia artificial ha aparecido en la escena de diversos sectores, sin que el ámbito jurídico haya sido la excepción. La conexión entre la violencia basada en género y la Inteligencia artificial se da de la mano de las necesidades crecientes de prevención de la primera, a través por ejemplo de la predicción de niveles de riesgo en la que la segunda ofrece importantes ventajas. Utilizando una metodología cualitativa deductiva con alcance descriptivo exploratorio, en la que se aplican métodos propios del derecho y las ciencias computacionales para analizar fuentes primarias, secundarias y estudio de casos de algoritmos y herramientas de evaluación de riesgo, (sin dejar de lado la referencia a herramientas de predicción tradicionales que no utilizan Inteligencia artificial), se arriba a resultados que apuntan a que los algoritmos y herramientas mencionadas evalúan y ponderan factores situacionales y disparadores, relacionados con el perpetrador, la víctima, y la relación familiar; variando en el valor asignado a cada

<sup>1</sup> Este artículo es producto del proyecto INV DER 3433 "Inteligencia Artificial aplicada a la predicción de la violencia de género" de la línea de investigación en derecho penal y justicia penal militar, del Grupo de Derecho Público, que se adelanta en el Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho de la Universidad Militar Nueva Granada, financiado por la Vicerrectoría de Investigaciones de la Universidad Militar Nueva Granada, dentro de la convocatoria de Proyectos de Investigación Científica vigencia 2021.

<sup>2</sup> Profesora e investigadora del grupo "Derecho Público" y de la línea "Derecho Penal y Justicia Militar" del Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho, Bogotá de la Universidad Militar Nueva Granada. Abogada por la Universidad del Rosario (Colombia). Especialista en Ciencias Penales y Ciencias Penales de la Universidad Externado (Colombia). Magister en derecho penal de la Universidad de Santo Tomás en convenio con la Universidad de Salamanca (Colombia). [marcela.roa@unimilitar.edu.co](mailto:marcela.roa@unimilitar.edu.co)

<sup>3</sup> Profesor e investigador del grupo "Derecho Público" y de la línea "Derecho Internacional, Derechos Humanos y Derechos Humanos Internacionales" del Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho, Bogotá de la Universidad Militar Nueva Granada. Abogado por la Universidad Militar Nueva Granada (Colombia). Especialista en Derechos Humanos y Derecho Internacional Humanitario aplicado a los Conflictos Armados de la Facultad de Derechos Humanos y Derecho Internacional Humanitario del Ejército Nacional (Colombia). Magister en Derecho Público Militar de la Universidad Militar Nueva Granada (Colombia). [jesus.sanabria@unimilitar.edu.co](mailto:jesus.sanabria@unimilitar.edu.co)

<sup>4</sup> Asistente de investigación de la Universidad Militar Nueva Granada. Abogada por la Universidad Militar Nueva Granada (Colombia). Especialista en Derechos Humanos y Derecho Internacional Humanitario de la Universidad Nacional de Colombia. [u0304087@unimilitar.edu.co](mailto:u0304087@unimilitar.edu.co)

uno de estos; en cuanto a las críticas se encuentran estandarizadas en la precisión y confiabilidad de la predicción.

**Palabras clave:** inteligencia artificial; violencia de género; feminicidio; predicción; niveles de riesgo.

---

#### ABSTRACT

Gender-based violence has been defined by the United Nations as any harmful act based on socially ascribed gender differences. Among its many manifestations, that violence in its maximum expression goes as far as femicide; a phenomenon that far from diminishing, has spread around the world.

On the other hand, artificial intelligence has appeared on the scene in various sectors, with the legal field being no exception. The connection between gender-based violence and artificial intelligence goes hand in hand with the growing prevention needs of the former, through, for example, the prediction of risk levels in which the latter offers important advantages.

Using a deductive qualitative methodology with an exploratory descriptive scope, in which methods from law and computer science are applied to analyze primary and secondary sources and case studies of algorithms and risk assessment tools (without leaving aside the reference to traditional prediction tools that do not use artificial intelligence), the results show that the algorithms and tools mentioned evaluate and weight situational factors and triggers, related to the perpetrator, the victim, and the family relationship; The criticisms are standardized in terms of the accuracy and reliability of the prediction.

**Keywords:** artificial intelligence; gender-based violence; femicide; prediction; risk levels.

---

#### INTRODUCCIÓN

La Violencia basada en género (en adelante VBG) ha sido definida por la ONU en su glosario de igualdad de género, como "un término genérico para cualquier acto perjudicial incurrido en contra de la voluntad de una persona, y que está basado en diferencias socialmente adjudicadas (género) entre mujeres y hombres" (ONU MUJERES, 2022, párr. 6)

Se trata de un problema de salud pública, que afecta a mujeres y niñas de todo el mundo. De acuerdo con la Organización Mundial de la Salud (en adelante OMS), "alrededor de una de cada tres (35%) mujeres en el mundo han sufrido violencia física y/o sexual de pareja o violencia sexual por terceros en algún momento de su vida" (OMS, 2017, párr 5).

La VBG adopta variadas formas y reviste diversos niveles de intensidad. El feminicidio es en muchas ocasiones la culminación de un largo ciclo de violencia en contra de la mujer, sin embargo, aún existen dificultades para determinar sus cifras reales, debido al bajo nivel de denuncias, revictimización, ceguera de

género de funcionarios encargados de la investigación y/o judicialización del delito.

La CEPAL, usando información oficial de quince (15) países de América Latina y tres (3) países del Caribe, ha señalado que 4.555 mujeres fueron víctimas de feminicidio o femicidio en 2019. Al sumar los datos de los cinco (5) países de la región que solo registran los feminicidios cometidos a manos de la pareja o ex pareja de la víctima (Barbados, Chile, Nicaragua, Puerto Rico y Suriname), se puede afirmar que el total de feminicidios ha sido de 4.640 mujeres para 2019 (ONU MUJERES, 2022, párr 9).

Para el caso colombiano, ONU mujeres señala que:

la violencia homicida contra las mujeres durante el quinquenio 2015 a 2019 dejó 5.013 víctimas, con una tasa promedio de 4,12 por cada 100.000 mujeres. Esto significa 1.484 víctimas menos que las registradas durante el quinquenio 2010-2014. En promedio, en este periodo fueron asesinadas 2,7 mujeres por día, la edad media estimada fue de 33,3 años. El escenario donde se registró el mayor número de casos fue la vivienda (35,77%), seguido de la vía pública (34,42%) (ONU MUJERES, 2022, párr 12).

Por su parte en el global database on violence against women, se indica que en Colombia el porcentaje de violencia física o violencia sexual a manos de la pareja, a lo largo de la vida de las mujeres, asciende al 33,3 % (UN Women, s.f., párr. 1).

A pesar de que los Estados intentan a través de las agencias de control, y autoridades en general, luchar contra este fenómeno, lo cierto es que las cifras demuestran que el crecimiento e intensificación del mismo no se detiene. En ese escenario, la posibilidad de establecer niveles de riesgo de feminicidio es una necesidad apremiante.

De otro lado, nos encontramos presenciando grandes avances en cuanto a la implementación de Inteligencia Artificial (en adelante IA), la cual, empleando técnicas como el Machine Learning (ML) es capaz, entre otras, de generar algoritmos de predicción que permitan establecer niveles de riesgo de violencia. De tal forma, en la lucha por la eliminación y persecución de crímenes violentos, se han desarrollado diversas herramientas de evaluación del riesgo las cuales, de acuerdo con (Ferguson & McLachlan, 2020) en su mayoría “son específicamente diseñadas para predecir la violencia y su escalada, no necesariamente homicidio”; sin embargo, pueden ya encontrarse algunas herramientas, que en efecto pretenden predecir el riesgo de homicidio íntimo, o aquel producido por la pareja, o expareja. También precisan los autores, que los factores medidos por estas herramientas examinan el riesgo de una escalada de violencia, revictimización o reincidencia, y solo a veces la letalidad (2020, pág, 2), (traducción libre).

Es así como se han desarrollado herramientas que buscan determinar el nivel de riesgo de feminicidio; dentro de las cuales algunas usan IA. En el caso español se encuentra VIOGÉN y SAS IBERIA además del modelo matemático de Erik Leal Enríquez, Certeza de Voz, Project Safe y BINDI (los cuales usan IA); y por otra parte, dentro de las herramientas de evaluación de riesgo (que no usan IA), se encuentran el ODARA y SARA.

En cuanto a la construcción de esas herramientas es pertinente analizar, cómo deben construirse, cómo se recoge la información, cómo deben escogerse y ponderarse las variables y cuáles deberían excluirse por tener un componente subjetivo que pueda llegar a generar sesgos. Una de las principales dificultades en el caso de algoritmos, es la denominada “black-box” o caja negra, término que ha sido utilizado para expresar el desconocimiento del funcionamiento interno del algoritmo. Si bien el programador humano establece los parámetros de funcionamiento, es el propio sistema el que los utiliza para la generación de información, lo cual impide en algunas ocasiones conocer el procesamiento de los datos que precede al resultado; ello implica uno de los elementos más cuestionados en este tipo de herramientas de aprendizaje automático.

Ahora bien, las herramientas que utilizan algoritmos inteligentes, son alimentadas con bases de datos que contienen información relacionada con la temática (por ejemplo la VBG); en un primer momento se realiza una clasificación de acuerdo a los parámetros establecidos por el diseñador humano, seguidamente se procede al análisis y aprendizaje, para llegar finalmente a un resultado de clasificación del nivel de riesgo, que puede estar dividido en tantas categorías como se establezca y requiera la arquitectura del prototipo.

Una de las principales características de estas herramientas que usan IA e implementan Machine Learning, es que, dentro de su programación, tienen un punto en el cual inician un proceso de aprendizaje, a lo anterior se suma que tienen la capacidad de analizar grandes niveles de información, lo que para un humano sería imposible de realizar con la misma eficacia y rapidez. Debido a lo anterior, a pesar de que en un inicio es el mismo diseñador humano el que establece los parámetros y reglas para el algoritmo, en cierto punto, éste utiliza este entrenamiento para encontrar relaciones entre los datos y generar resultados, lo que en palabras sencillas significa que el algoritmo se separa de los parámetros del creador, ganando cierta independencia en el proceso.

De acuerdo con esta panorámica presentada, la pregunta de investigación que surge es ¿Cuáles son las dinámicas existentes entre la predicción de VBG y riesgo de feminicidio y las herramientas de predicción, en especial, aquellas que utilizan Inteligencia Artificial?

Para dar respuesta a la misma se utiliza una metodología cualitativa deductiva con alcance descriptivo exploratorio, en la que se aplican métodos propios del

derecho y las ciencias computacionales para analizar fuentes primarias, secundarias y estudio de casos de algoritmos y herramientas de evaluación de riesgo, sin dejar de lado la referencia a herramientas de predicción tradicionales que no utilizan IA. Los resultados apuntan a que se evalúan y ponderan factores situacionales y disparadores, algunos relacionados con el perpetrador, la víctima, y la relación familiar (tipo de violencia anterior, amenazas de homicidio etc.), las diversas herramientas parecen utilizar un grupo similar de factores, variando en el valor asignado a cada uno de estos; incluso las críticas se encuentran estandarizadas en la predicción sobre la precisión y confiabilidad.

Por último, los resultados se presentarán de la siguiente forma: primero se contextualizará la VBG, el feminicidio como máxima expresión de aquella, la violación de derechos a causa de estos crímenes; seguidamente se describirán los sistemas de predicción de riesgo de VBG, dentro de las cuales están: SARA, VIOGÉN, ODARA, el modelo matemático de Erik-Leal Enríquez, certeza de voz, entre otros. Posteriormente se mencionarán las variables que son ponderadas por estos sistemas, y finalmente se describirán las ventajas y desventajas de dichas herramientas; para dejar abierta la posibilidad de incorporar factores novedosos derivados de la entrada en el escenario de las neurociencias y en particular de la neurocriminología.

### **1. La VBG y el Feminicidio como una violación de derechos contra niñas y mujeres**

La VBG ha permeado de formas devastadoras las esferas cotidianas en ámbitos laborales, académicos y familiares, al punto de convertirse en un problema de salud pública. Según lo afirmado por la OMS, afecta a mujeres de todas las etnias, edades, orientación política, orientación sexual, en razón de su género. Los intentos por definir el fenómeno de la VBG parecen encontrar puntos en común con la propuesta de ONU MUJERES, la cual precisa que se refiere a los actos dañinos dirigidos contra una persona o un grupo de personas en razón de su género. Tiene su origen en la desigualdad de género, el abuso de poder y la existencia de normas dañinas. El término se utiliza principalmente para subrayar el hecho de que las diferencias estructurales de poder basadas en el género colocan a las mujeres y niñas en situación de riesgo frente a múltiples formas de violencia (s.f, párr 2).

Por otra parte, Bolívar & Erazo la definen como un fenómeno de carácter estructural, social, político y relacional que constituye una violación a los derechos humanos [...] lo cual afecta principalmente a las mujeres, no obstante, no excluye a personas con otras identidades de género (2019, pág. 183).

En el mismo sentido, la Organización de Estados Americanos (en adelante OEA) (1994), ha señalado que la violencia dirigida hacia la mujer es “cualquier acción o conducta [...] que cause muerte, daño o sufrimiento físico, sexual o psicológico a la mujer, tanto en el ámbito público como en el privado” (pág.82). El culmen

de dicha violencia es el feminicidio, un fenómeno que sucede como consecuencia de ciclos de violencia en contra de la mujer por el hecho de serlo, y cuyas cifras siguen creciendo a nivel global. El feminicidio, inicialmente nombrado así por Leonore Walker, ha sido estudiado en sus diferentes variantes. En ese sentido ONU mujeres clasifica el fenómeno del femicidio entre otras, tomando en cuenta la relación existente entre la víctima y el agresor, así: i) feminicidio de pareja íntima, ii) feminicidio de familiares, iii) feminicidio por otros conocidos y iv) feminicidio de extraños (2022, párr.1).

En cuanto a la magnitud del feminicidio, el Observatorio de Igualdad de Género de América Latina y el Caribe, refiere que:

En América Latina las tasas más altas por cada 100.000 mujeres de feminicidios o femicidios en 2020 corresponden a Honduras, República Dominicana y El Salvador. Los nueve (9) países que informaron una disminución en las tasas de feminicidio en comparación con el año 2019, son Bolivia, Brasil, Colombia, El Salvador, Guatemala, Honduras, Paraguay, República Dominicana y Uruguay. Argentina y México mantuvieron las mismas tasas de feminicidio que en 2019, mientras que tres (3) países (Ecuador, Costa Rica y Panamá) registraron un aumento en comparación con el año anterior. (2019, párr. 2).

En el mismo panorama, según un informe de la Organización Mundial de la Salud en conjunto con la Organización Panamericana de la Salud (2013) en Estados Unidos las mujeres embarazadas pueden tener un mayor riesgo de feminicidio cometido por su pareja según un examen de registro policial realizado en once (11) ciudades de Estados Unidos (pág. 2).

Ante tal situación, se han desarrollado diversos cuerpos normativos, dentro de los cuales se encuentran, por ejemplo, el Violence Against Women Act, VAWA (Estados Unidos), Ley 1761 de 2015 (Colombia), Ley del Femicidio (Brasil), Ley Especial Integral para una Vida Libre de Violencia para las Mujeres Integral contra la Violencia Hacia las Mujeres (Guatemala).

A nivel internacional los principales instrumentos son la Convención para la Eliminación de todas las formas de Discriminación contra la Mujer (CEDAW), de Naciones Unidas, la Convención Interamericana para Prevenir, Sancionar y Erradicar la Violencia contra la Mujer (Convención de Belém do Pará) de la OEA, Convenio del Consejo de Europa sobre prevención y lucha contra la violencia contra las mujeres y la violencia doméstica, del Consejo de Europa.

No obstante, la existencia de normatividad a nivel nacional e internacional, y la creciente concientización en torno a la gravedad de este fenómeno, queda en evidencia que la mujer sigue siendo víctima de diversos tipos de violencia, y en los casos más graves, de feminicidio. Esta realidad conlleva la transgresión de diversos derechos humanos, tales como la vida, la dignidad, la integridad física y moral, la igualdad, la seguridad, la libertad, la autonomía, y en últimas a la garantía de una vida libre de violencias.

Hasta este punto, puede concluirse que las solas medidas normativas no resultan suficientes ni eficientes en la disminución de la VBG, y es allí en donde la IA tiene un papel relevante ya que a través de algunos algoritmos y herramientas de evaluación de riesgo se ha planteado la posibilidad de determinar o predecir el nivel de riesgo de violencia íntima y feminicidio, lo que permitiría a su vez a las autoridades, anticiparse a la decisión criminal del feminicida, y establecer medidas de protección oportunas que eviten la materialización de este delito.

## **2. Sistemas de Predicción de Riesgo de VBG**

La IA es una disciplina que posee un espectro de aplicaciones en casi cualquier área del conocimiento; ella brinda la capacidad a una máquina de resolver tareas de manufactura y toma de decisiones, imitando la capacidad del ser humano para ejecutarlas (IBM, 2020). Puede definirse también “como la habilidad de un sistema para interpretar correctamente datos externos, aprender de dichos datos y utilizar el aprendizaje obtenido para alcanzar objetivos específicos a través de una adaptación flexible” (Kaplan y Haenlein, 2019, pág. 17).

Ahora bien, en cuanto a su aplicación, la IA es clasificada mediante dos categorías, por una parte, la Inteligencia Artificial Débil (IAD) y de otra, la Inteligencia Artificial Fuerte (IAF). La primera está enfocada en la ejecución de tareas únicas o limitadas, como el manejo de vehículos autónomos o los asistentes de acompañamiento por voz como Alexa, Siri o Cortana (IBM, 2020).

Por otro lado, la segunda implica la capacidad de una máquina para tener conciencia y aprender como lo haría un ser humano. Sin embargo, este tipo de IA es meramente especulativa y aún se plantea su posible desarrollo (Cheisvili, 2021).

Por su parte, una de las técnicas derivadas de la IA es el Machine Learning (ML), [...] la cual permite la toma de decisiones a partir del análisis de grandes conjuntos de datos específicos, por medio de modelos que son diseñados y entrenados para la detección de patrones, variaciones estadísticas y valores entre dichos datos, con el fin de mejorar su rendimiento a partir de la experiencia que adquieren con su proceso de aprendizaje [...](Surden, 2018, pág. 1311). En un algoritmo de ML, los datos con los que se entrena el sistema son etiquetados y el resultado final puede separarse según el tipo de problema que se busque predecir, ya sea un problema de clasificación (asignar datos de prueba en categorías) o de regresión (medir la relación entre una variable dependiente y otra independiente) (IBM, 2021).

Para ello, resulta importante mencionar que un algoritmo que hace uso de la técnica de minería de datos funciona como un clasificador de aprendizaje automático, es decir, que extrae información que no es detectada a simple vista

en volúmenes extensos de datos, reconociendo patrones entre estos y descubriendo relaciones o excepciones entre variables, que puedan ser útiles para la toma de decisiones.

La técnica de ML ha sido utilizada principalmente en análisis de finanzas, visión por computador, procesamiento del lenguaje, en la industria de energía y automotriz (Mathworks, s.f.), recientemente ha empezado a utilizarse en el ámbito de la administración de justicia, donde se ha puesto a prueba su efectividad para predecir el riesgo de litigio sobre un conjunto de datos establecidos (Lee et al, 2019), también al realizar análisis semántico de textos legales (Yablon y Roos, 2012; Liu y Chen, 2017,), y en SPRV, los cuales hacen uso de técnicas de ML para predecir una conducta violenta en escenarios domésticos (Berk et al, 2016), así como en problemáticas de violencia de género (González-Prieto et al, 2021; Bello et al, 2020; Rodríguez-Rodríguez et al, 2020;), entre otros.

La aplicación de la IA se ha materializado también en el campo de los Sistemas de Predicción de Riesgo de Violencia (SPRV), en los que, mediante un procedimiento de aprendizaje supervisado, partiendo de un conjunto de datos se interpreta y predice una conducta violenta relacionada con una población o tipo de violencia específica. En palabras sencillas, se obtiene como resultado una predicción a partir de unos datos de entrada (González-Prieto et al, 2021).

A partir de las clasificaciones y categorías de IA anteriormente mencionadas, se procederá a describir brevemente algunos algoritmos y algunas herramientas de evaluación de riesgo, empezando por aquellas que no utilizan IA, pero que constituyen el antecedente de aquellas, y que son todavía ampliamente utilizadas a nivel mundial.

#### a. SARA

El Spousal Abuse Risk Assessment (en adelante SARA) es una herramienta de evaluación que permite determinar el nivel de riesgo de feminicidio, riesgo que puede ser físico o sexual. Al respecto, Pueyo, López & Álvarez (2008, pág. 115-116) precisan que el conocimiento de la técnica de la valoración del riesgo para la predicción de la violencia nos dice que nunca podremos saber si una persona realizará un determinado acto violento en el futuro, sólo podremos estimar la probabilidad de que en determinadas ocasiones y condiciones (en un entorno familiar, escolar, etc.) y para un intervalo temporal limitado (semanas o meses) aparezca un acto violento.

SARA utiliza veinte (20) factores de riesgo relacionados con la violencia contra la pareja con el fin de predecir la probabilidad de que el agresor pueda reincidir en la conducta violenta después de realizada dicha valoración. Los factores de riesgo mencionados se agrupan en cinco ítems, así:



**Tabla 1**

<p><b>1. Historial delictivo</b></p>	<p>Se analiza y recolecta información de sucesos delictivos, los cuales no necesariamente están relacionados con delitos de violencia intrafamiliar. En este ítem se incluye, “Violencia anterior contra los familiares”, “Violencia anterior contra desconocidos o contra conocidos no-familiares y la “Violación de la libertad condicional u otras medidas judiciales similares”</p>
<p><b>2. Ajuste psicosocial.</b></p>	<p>Relacionado con desordenes psicosociales dentro de los cuales se pueden encontrar, “Consumo/Abuso reciente de drogas”, “Ideas/intentos de suicidio y/o homicidios recientes, “Trastorno de personalidad con ira, impulsividad e inestabilidad conductual”, entre otros.</p>
<p><b>3. Historia de violencia en pareja</b></p>	<p>En este ítem se encuentran algunos sub-ítems: “Violencia física anterior”, “Violencia sexual y/o ataque de celos en el pasado”, “uso de armas y/o amenazas de muerte creíbles”, “incremento reciente en la frecuencia o gravedad de las agresiones”.</p>
<p><b>4. Delito/agresión actual (que motiva la valoración).</b></p>	<p>Se relaciona con la agresión más reciente que se haya materializado y que incluya, “Violencia sexual grave”, “uso de armas y/o amenazas de muerte creíbles” y “violación o incumplimiento de las órdenes de alejamiento”.</p>
<p><b>5. Otras consideraciones</b></p>	<p>Este apartado está disponible para que el evaluador consigne aquellas situaciones que considera como relevantes para el caso en cuestión y que puedan aportar al grado de riesgo de la violencia de pareja, por ejemplo, sadismo sexual o mutilación de las parejas sexuales.</p>

**Fuente:** Elaboración propia, información tomada de (Pueyo, López, & Álvarez, 2008, pág. 116)

Ahora bien, en relación con el funcionamiento de esta evaluación de riesgo, Pueyo, López y Álvarez señalan que,

la tarea de resumir la valoración final del riesgo de violencia mediante el uso de esta “guía” se hace de forma no-reglada ni ponderada cuantitativamente, es decir sin seguir un algoritmo preciso de decisión, por lo que implica la presencia del juicio del evaluador. Generalmente se hace teniendo en cuenta el número de ítems que están presentes en la valoración y de aquellos ítems críticos destacados en la misma. La valoración final se resume en cuatro niveles que son: bajo, moderado, elevado e inminente (Pueyo et al, 2008, pág. 117).

El resultado obtenido en la valoración de riesgo es usado en diferentes escenarios, tales como procesos policiales, judiciales, de protección y seguridad, servicios dirigidos a las víctimas, entre otros. Es claro, que aquí no hay presencia de IA, y se parte de un ejercicio de valoración humano de los factores contemplados en la guía, sin que exista un criterio uniforme. No obstante lo anterior, SARA es una de las herramientas más antiguas, y en las que por primera vez se estableció ese catálogo de factores, que ha sido el insumo para posteriores desarrollos.

## b. VIOGÉN (ESPAÑA)

El llamado algoritmo VIOGÉN, es un

Sistema de Seguimiento Integral en los casos de Violencia de Género, de la Secretaría de Estado de Seguridad del Ministerio del Interior, el cual se puso en funcionamiento en el año 2007, en cumplimiento de lo establecido en la Ley Orgánica 1/2004, de 28 de diciembre sobre las Medidas de Protección Integral contra la Violencia de Género. Dentro de sus objetivos se mencionan el aglutinar las diferentes instituciones públicas que tienen competencias en materia de violencia de género, integrar toda la información de interés que se estime necesaria, hacer predicción del riesgo, una vez establecido el nivel de riesgo, realizar seguimiento y protección a las víctimas en todo el territorio nacional, entre otros (Ministerio del Interior, 2022, párr 1-2).

Como muchos de los sistemas que pretenden predecir y en consecuencia prevenir la violencia de género y el feminicidio, resalta dentro de los propósitos, el mejorar la efectividad en la protección de las víctimas, a través de la emisión de alertas tempranas, seguimiento del riesgo, y selección de las medidas de protección necesarias para las víctimas. (Zurita, 2013) citado en (González Álvarez et al, 2018, pág. 31).

Es así como al vincular a las diversas autoridades que a nivel nacional intervienen en el tratamiento, prevención y sanción de la violencia de género, se pretende también, facilitar la evaluación del riesgo que, a través de la integración de la información, permite monitorear las variaciones de tal riesgo a través de alarmas y notificaciones automáticas, e inclusive diseñar un plan personalizado de seguridad de la víctima (Zurita, 2013, como se cita en González Álvarez et al, 2018, pág. 31)

Aunque VIOGÉN parte del diligenciamiento de dos formatos clave, que son el de valoración policial del riesgo (VPR) y el de valoración policial de evolución del riesgo (VPER), a la vez cuenta con un sistema informático on line, al que acceden las múltiples agencias que intervienen en la prevención, investigación y sanción de la VBG (González Álvarez et al, 2018, págs. 43-44).

Es interesante que en la construcción de los textos finales de los formularios se contó con la participación de profesionales de diversas profesiones, incluyendo psicólogos, funcionarios de policía y académicos; quienes acordaron los indicadores que debían incluirse, y entre los que se encuentran, por ejemplo, factores de vulnerabilidad de la víctima (discapacidad, ser menor de edad o tener menores a su cargo), factores de peligrosidad del agresor y violencia en entornos escolares, entre otros (González Álvarez et al, 2018, págs. 43-44)

Respecto de los dos formatos clave anteriormente mencionados, es necesario indicar que el VPR se compone de treinta y nueve (39) indicadores, agrupados en cuatro (4) dimensiones temáticas, con doce (12) factores (González Álvarez

et al, 2018, pág. 57), para mayor claridad, esas dimensiones y su contenido pueden esquematizarse así:

**Tabla 2**

<b>Primera Dimensión</b>	Tiene en cuenta, la gravedad del hecho e historia de violencia de la pareja (se explora si ha habido amenazas, uso de armas, por ejemplo)
<b>Segunda Dimensión</b>	Se enfoca en el agresor (celos, acoso, comportamientos antisociales, o signos psicopatológicos). Se indaga, por ejemplo, si el agresor tiene antecedentes, intentos de suicidio, trastornos psicológicos o psiquiátricos.
<b>Tercera Dimensión</b>	Se enfoca en características de la víctima que la ponen en posición de vulnerabilidad, al igual que factores situacionales que puedan detonar violencia (discapacidad, gestación, intentos de suicidio, adicción, dependencia económica)
<b>Cuarta Dimensión</b>	Percepción de la víctima sobre su situación (conciencia de su situación).

**Fuente:** Elaboración propia, información tomada de González Álvarez et al. (2018, pág. 57).

Por su parte, el VPER, ha sido definido como un:

instrumento semiactuarial, que utiliza indicadores estáticos y dinámicos (de riesgo y de protección) [...] tiene funcionamiento totalmente autónomo, con un algoritmo propio y es capaz de realizar pronósticos mediante un conjunto mixto (riesgo-protección) de indicadores que permiten evaluar la probabilidad de reincidencia y al mismo tiempo monitorizar los cambios que se producen a lo largo del tiempo (González Álvarez et al, 2018, pág. 61).

Este segundo formulario consta de cuarenta y tres (43) indicadores, dentro de los que treinta y cuatro (34) son de riesgo y nueve (9) de protección, los cuales se encuentran agrupados en las mismas dimensiones del VPR. Dentro de los factores se incluyen, por ejemplo, si el agresor se ha fugado o se encuentra en paradero desconocido o si la víctima dificulta a la policía el poder tomar acciones para su protección (se retracta, reanuda la convivencia con el agresor); de igual forma, también se incluyen las evidencias de comportamientos “responsables del agresor”, tales como, alejarse de la víctima, actitud pacífica, mostrar arrepentimiento, acogerse a programas de ayuda, etc.

Es interesante, que de acuerdo con el nivel de riesgo establecido (no apreciado, bajo, medio, alto, extremo), se disponen medidas obligatorias y complementarias que buscan proteger a la víctima. En este aspecto, VIOGÉN cuenta con un elemento denominado plan de seguridad personalizado, y un segundo elemento complementario, es el denominado “caso”, al interior del sistema. Respecto del primero, se trata de un plan que se genera y aplica tanto para las mujeres víctimas como para los menores a su cargo, y responde a la idea de implicar a la víctima en todo el proceso de construcción de este,

ajustando las medidas según las necesidades y el nivel de riesgo establecido para dicha víctima.

En relación con el segundo, “un caso contiene toda la información que relaciona a una víctima con un único agresor, de manera que, si una mujer a lo largo del tiempo es víctima de violencia de género con más de un agresor, existirá un caso distinto por cada uno de los diferentes agresores. Ocurre lo mismo cuando un agresor maltrata a diferentes mujeres” (González Álvarez et al, 2018, pág. 71).

Debe también señalarse que, existiendo conexión entre las diversas agencias del estado, los usuarios del sistema que pueden incluir información y modificar datos, incluyen, entre otros, a jueces y fiscales, directores de penitenciarías, Instituto de Medicina Legal, servicios asistenciales y servicios sociales. Todos estos agentes, se identifican con un usuario, contraseña y tienen perfiles diferenciados para cada entidad. (González Álvarez et al, 2018, pág. 75). El sistema conecta diversos tipos de información y busca que se puedan intercambiar datos con facilidad, por lo que se integra a través de servicios web, con información de la policía, sistema penitenciario y sistema judicial.

#### c. SAS IBERIA

Luego de la implementación del VIOGÉN; en el año 2021 se incorporó un software adicional, que pretende optimizar el trabajo de la versión inicial de 2007; en particular, se pretendía “actualizar más eficaz y rápidamente, optimizando la ponderación de los actuales indicadores de riesgo de reincidencia, e identificando nuevas variables en periodos de tiempo más cortos” (Flores, 2021, párr 2).

La empresa creadora de esta herramienta ha indicado, que con la incorporación del SAS, se pretende “acelerar el proceso de clasificación del riesgo de reincidencia de agresiones y ser capaz de atender a los casos más graves, más rápido y con los recursos necesarios” (SAS, 2020, párr 1).

La herramienta SAS es calificada como un agente digital, es decir que funciona de forma automática, ahorrando el uso de recurso humano, al permitir analizar grandes cantidades de datos mediante la automatización de ese proceso. Adicionalmente, la empresa desarrolladora asegura que este nuevo nivel de análisis en los datos permite “hacer una gestión más proactiva de los casos de reincidencia y, ayuda a los cuerpos de seguridad a hacer una mejor asignación de recursos de acuerdo con el grado de riesgo de la víctima” (SAS, 2020, párr 3). El objetivo de esta herramienta es hacer análisis “casi en tiempo real” (en palabras de SAS), tanto de la información contenida en las denuncias como de todos los demás factores que pondera VIOGÉN logrando así optimizar recursos y tiempo. SAS promete que se pueda realizar

un análisis con más de cincuenta (50) indicadores, de manera automatizada en menos de una hora gracias a la nueva tecnología. Esta agilidad ganada en el proceso de clasificación del riesgo permite a las fuerzas de seguridad tener un mapa actualizado de los casos de riesgo, casi en tiempo real, lo que aumenta la eficiencia a la hora de asignar recursos para la protección de las víctimas. Las tecnologías de inteligencia artificial y machine learning de SAS son capaces de distinguir aquellos casos donde se va a producir algún tipo de reincidencia, predecir cuándo y cuántas veces va a ocurrir (SAS, 2020, párr 5).

Respecto de esa herramienta, la fundación mujeres<sup>5</sup> precisa que con ella se

introduce el concepto de smart policing en la gestión de la lucha contra la violencia de género, se combinan datos de diversos entornos, se mejora la tecnología con la son analizados y se valora el riesgo que existe para las víctimas, y se posibilita una toma de decisiones más eficiente sobre los casos que deben seguir activos y los recursos que deben ser asignados (Flores, 2021, párr 4).

Así las cosas, este complemento del ya implementado VIOGÉN, ofrece la mejora respecto del Protocolo de Valoración del riesgo, junto con el punto más importante dentro de la actualización que es el análisis en tiempo real. Todo esto, buscando favorecer la oportunidad, en la toma de decisiones que se requieran para la protección de las víctimas. Un dato relevante es que en entrevista con el diario El País, el director general de SAS IBERIA, introdujo el movimiento “Data4Good, el cual se trata de una iniciativa que fomenta el uso de datos para resolver problemas relacionados con la pobreza, salud, derechos humanos, educación y el medio ambiente” (Granados, 2021, párr 2). Al respecto debe indicarse que, en materia de prevención de la VBG y del feminicidio, es imperativo que se abra el espacio para que la tecnología, y en especial la IA para ofrecer verdaderas soluciones, que redunden en mayores niveles de igualdad y vida libre de violencias para las mujeres.

## ODARA

ODARA se define como una herramienta actuarial <sup>6</sup> sencilla de evaluación de riesgo de violencia doméstica, la cual es utilizada por agentes de policía,

---

<sup>5</sup> Se puede ver información en <https://observatorioviolencia.org/quienes-somos/>

<sup>6</sup> Una práctica actuarial de toma de decisiones en la evaluación de riesgos se refiere al uso de métodos matemáticos y estadísticos aplicados a una serie de variables para predecir el nivel de ocurrencia de un evento. Aunque son comúnmente usados en áreas como finanzas, pensiones, seguros o inversiones; su adaptación en el campo judicial intenta predecir la probabilidad de reincidencia del delincuente teniendo en cuenta diversos factores que puedan incidir en el resultado. Un modelo o herramienta actuarial tiene la función de conceptualizar el riesgo de violencia en términos de probabilidad, están diseñados únicamente para prever futuras agresiones basados en cuestionarios o ítems que se evalúan en escalas donde, dependiendo del resultado, generan suposiciones relativas a la gravedad de los acontecimientos.

trabajadores de servicios, tribunales, etc., (Ulmer, 2015, pág. 9) Respecto de los ítems valorados por este instrumento, se relacionan los siguientes:

1. Violencia en relaciones anteriores
2. Violencia previa sin relación
3. Sentencia correccional previa de 30 días o más
4. Incumplimiento de liberaciones condicionales anteriores
5. Índice profiriendo amenazas
6. Índice de internamiento ilegal
7. Víctima temerosa de reincidencia
8. Dos o más hijos en la familia
9. La víctima tiene un hijo biológico de una relación anterior
10. Historia de violencia fuera del hogar.
11. Historial de abuso de sustancias
12. Víctima embarazada en el momento de la agresión
13. La víctima tiene barreras hacia el apoyo (Risk Management Authority, 2019).

Respecto de los ítems o factores utilizados por ODARA, es preciso indicar que “el contenido del elemento es predominantemente de naturaleza estática e incluye antecedentes penales, detalles de los delitos índice y datos demográficos de la víctima y la familia”. (Hegel, Pelletier, & Olver, 2021, pág. 3). A pesar de que algunos estudios habían señalado que ODARA tenía una precisión moderada en la predicción, respecto de una variante nueva en la que se incluyeron catorce (14) variables relacionadas con el check list de psicopatía se encontró que la precisión en la predicción de la reincidencia para agresión conyugal era mucho más alta (Hegel, Pelletier, & Olver, 2021, pág. 3).

Esta herramienta ha sido utilizada en diversas ciudades y condados de los Estados Unidos, y se encuentra que, en algunos de ellos, se han establecido sugerencias acerca de cómo establecer el contenido de cada factor, a continuación, puede verse, de forma ejemplificativa, como el Condado de Mayne ha sugerido tal contenido.

**Tabla 3**

FACTOR	CONTENIDO
<b>1. Violencia en relaciones anteriores</b>	Consiste en cualquier incidente en el que la persona evaluada haya agredido a su pareja actual o expareja íntima y/o hijo de esa víctima, o pareja con la que esté teniendo citas, y que haya sido registrado en un informe policial o antecedentes penales.
<b>2. Violencia previa por fuera de relaciones</b>	Se tiene en cuenta cualquier incidente en el que la persona evaluada haya agredido a otra con la que no tenga ninguna relación íntima actual o anterior ni que se trate de la pareja actual o el hijo de esa persona, y que consta en un informe policial o antecedentes penales.
<b>3. Previa sentencia correccional de 30 días o más</b>	Este factor implica que respecto del individuo evaluado existe una decisión final expedida por un tribunal por un delito (la sanción debe haber sido de por lo menos 30 días)

<b>4. Incumplimiento de libertad condicional previa</b>	Se trata de violaciones a la libertad condicional, sin importar si se produjo o no arresto o informe policial al respecto (incluye por ejemplo el no presentarse a la Corte o contactar a una persona existiendo una prohibición judicial al respecto)
<b>5. Amenazas de daño o de muerte</b>	En este factor debe tenerse en cuenta cualquier amenaza de causar la muerte o daño físico, realizada a cualquier persona. La mencionada amenaza puede expresarse con palabras o con gestos evidentes. No es necesario que la amenaza se materialice.
<b>6. Confinamiento de la víctima</b>	Consiste en cualquier acto realizado para evitar que la víctima (pareja actual o anterior del agresor) abandone el lugar del ataque.
<b>7. Preocupación de la víctima por una futura agresión</b>	Se relaciona con manifestaciones de la víctima en las que expresa preocupación o miedo, por la certeza de que el evaluado, va a agredirla a ella o a su (s) hijo (s) en el futuro
<b>8. Más de un hijo</b>	Implica que la víctima, y el evaluado, tienen más de un hijo. Deben contarse los hijos que tengan de uniones anteriores, o los que tengan en común
<b>9. Hijo biológico de la víctima de una pareja anterior</b>	Se tienen en cuenta los hijos de la víctima, con un sujeto diferente al evaluado
<b>10. Incidente violento anterior contra una pareja no doméstica</b>	En este factor deben contarse todos los incidentes en los que el sujeto evaluado ha agredido a una persona diferente de su pareja actual o exparejas (ataque entendido como contacto físico o el intento de usar algún arma, o amenaza de causar daño esgrimiendo un arma)
<b>11. Dos o más indicadores de abuso de sustancias</b>	Este factor implica que el individuo evaluado ha sido acusado previamente por delitos relacionados con el uso de alcohol o sustancias (Ej. consumo de alcohol o drogas antes o durante el incidente, problemas con alcohol o drogas que hayan influido en la vida del evaluado después de cumplidos los 18 años)
<b>12. Agresión a la víctima índice cuando estaba embarazada</b>	Este factor se presenta cuando la agresión se dio contra una víctima que estaba embarazada, o cuando el evaluado haya agredido a la misma víctima en una ocasión diferente y anterior durante el embarazo actual o durante un embarazo anterior.
<b>13. Barreras para el apoyo a la víctima</b>	Cuando uno o más de las siguientes barreras estaban presentes en el momento del incidente índice. Ej. sin acceso a teléfono o vehículo, transporte público o dinero para el transporte, vivienda en un sitio aislado geográficamente, consumo o historial de abuso de drogas o alcohol por parte de la víctima

**Fuente:** Elaboración propia, información tomada de Maine Guidelines & General Scoring Criteria 2019, The Maine Coalition to End Domestic Violence [MECDV] y Maine Cumberland County

Para evaluar los factores recién mencionados, debe “calificarse cada uno de los trece (13) ítems con un puntaje de uno (1) si hay evidencia que el ítem está presente o cero (0) si no hay evidencia de que el ítem esté presente” (The Maine Coalition to end domestic violence and Violence Intervention Partnership of Maine Cumberland County, 2014, pág. 2) (traducción libre).

También se da la opción de calificar el factor con un signo de interrogación, cuando éste sea poco claro o ambiguo. El puntaje total será el resultante de la suma de los puntajes asignado a cada ítem, (The Maine Coalition to end domestic violence and Violence Intervention Partnership of Maine Cumberland County, 2014, pág. 10.11), (traducción libre). Respecto del recaudo de la información para ponderar los factores recién mencionados, ella proviene entre otras, de “entrevistas a las víctimas, revisión de historia criminal”. (Maine Guidelines & General Scoring Criteria - ODARA, pág. 3)

La calificación de los factores arroja un resultado en el que los puntajes altos son indicativos de que el sospechoso es más propenso a cometer otros delitos, en un periodo corto de tiempo y causar daños, que un individuo que haya obtenido puntajes bajos (Risk Management Authority, 2019).

Una de las ventajas que se ha señalado respecto de ODARA, es que se trata de una herramienta actuarial, y los estudios han demostrado que aquellas que utilizan estos métodos actuariales, gozan de una mayor precisión.

El puntaje ODARA está relacionado con una tabla actuarial que indica el riesgo asociado con cada puntaje, en términos de probabilidad (probabilidad de reincidencia de VPI en un período de tiempo determinado) y rango percentil (cómo se compara un individuo con otros en términos de riesgo de VPI (Radatz & Hilton, 2021, pág. 3), (traducción libre)

Sin embargo, ODARA no ha estado absolutamente libre de críticas, ya que algunos estudios han afirmado que su uso en Canadá puede haber estado sesgado respecto de las poblaciones indígenas, las cuales parecen estar sobrerrepresentadas en las muestras, lo que conllevaría a fallas o sesgos en la predicción. Por su parte en relación con la validación cruzada, Radatz & Hilton resaltan la importancia de validación de las herramientas con muestras diversas obtenidas, lo anterior, sin perder de vista las diferencias que pueden darse entre jurisdicciones o regiones, poblaciones con diversas características individuales, diferencias en la administración de justicia deben ser tenidas en cuenta ya que una herramienta diseñada para una población y contexto específicos puede no ser adecuada en otro (Radatz & Hilton, 2021).

Uno de los temas novedosos respecto de ODARA, es que se ha indicado que puede ser utilizado también para evaluar mujeres agresoras, lo cual se alinea con la noción precisa de VBG, la cual, sin duda, puede ser ejercida por parte de hombres o mujeres, siempre y cuando ella esté motivada por razones de género. En ese sentido,

investigaciones recientes han validado el uso de ODARA para ofensoras mujeres. Las mujeres con puntajes altos en ODARA son más proclives a reincidir, más que aquellas mujeres con puntajes bajos. En todo caso, no pueden compararse los puntajes de mujeres agresoras con los puntajes



de hombres agresores, porque en general la reincidencia de las mujeres es mucho más baja que la de los hombres; las mujeres reinciden a tasas mucho más bajas de lo esperado, basándose en lo esperado para agresores hombres (The Maine Coalition to end domestic violence y Violence Intervention Partnership of Maine Cumberland County, 2014, pág. 13), (traducción libre).

### c. El modelo matemático de Erik-Leal Enríquez

Erik Leal Enríquez, ha construido un modelo matemático, que al igual que otras herramientas tiene “como objetivo prevenir futuras lesiones relacionadas con IPV<sup>7</sup> u otros tipos de violencia al garantizar una intervención temprana para interrumpir el ciclo de violencia” (Leal-Enríquez, 2018, pág. 1061).

El creador precisa que, dentro de las herramientas diseñadas para medir la violencia en contra de la mujer, muchos de los instrumentos utilizan herramientas estadísticas para identificar la etapa y tipo de violencia que sufre la víctima de violencia de pareja, y se incluyen factores que influyen en la violencia; factores, que normalmente se relacionan con la frecuencia de los actos violentos, y la gravedad de los mismos (Leal-Enríquez, 2018, págs. 1052-1053)

Leal-Enríquez establece que el modelo, parte de la hipótesis de que la violencia acumulada puede ser “modelada”, analizando los meses previos; y aquí lo novedoso de esta propuesta, es que se indica que el factor clave para determinar la continuación o reducción de la violencia, es el probable estado de autocontrol del perpetrador (Leal-Enríquez, 2018, pág. 1053). A pesar de que este modelo sin duda incorpora un factor novedoso, que no ha sido medido o ponderado en otras herramientas, se señalan como limitaciones del mismo, que no toma en cuenta las edades de hombres y mujeres, ni estratos, factores culturales, que eventualmente podrían estar correlacionados con la pérdida de autocontrol (Leal-Enríquez, 2018, pág. 1060), lo que deja en evidencia que no habría algún tipo de sesgo en los resultados ponderados que brindaría el modelo matemático. A pesar de que este modelo sin duda incorpora un factor novedoso, que no ha sido medido o ponderado en otras herramientas, se señalan como limitaciones de este, que no toma en cuenta las edades de hombres y mujeres, ni estratos, factores culturales, que eventualmente podrían estar correlacionados con la pérdida de autocontrol (Leal-Enríquez, 2018, pág. 1060)

A pesar de tratarse de un modelo altamente complejo, se pueden establecer de manera general cómo funciona. De acuerdo con Leal-Enríquez, el modelo genera posibles escenarios de violencia por medio de simulaciones computacionales, en diversos niveles de severidad, fundamentalmente en atención al indicador de pérdida de autocontrol. Se tiene en cuenta el indicador

---

<sup>7</sup> Intimate Partner Violence (violencia de pareja/violencia íntima)

de prevalencia de la violencia en cierto periodo de tiempo, que esa prevalencia puede cambiar de mes a mes, la prevalencia de la violencia permite también calcular el estado probable de la pérdida de control por parte del perpetrador (Leal-Enríquez, 2018, pág. 1060)

Por otra parte, a pesar de tratarse de un modelo altamente complejo, se puede establecer de manera general cómo funciona. Su propio creador explica que el modelo genera posibles escenarios de violencia por medio de simulaciones computacionales, en diversos niveles de severidad; fundamentalmente en atención al indicador de pérdida de autocontrol. Se tiene en cuenta el indicador de prevalencia de la violencia en cierto periodo de tiempo, sin perder de vista que esa prevalencia puede cambiar de mes a mes, y que la prevalencia de la violencia permite también calcular el estado probable de la pérdida de control por parte del perpetrador (Leal-Enríquez, 2018, pág. 1060)

El modelo utiliza algo denominado el indicador dinámico, el cual puede ayudar a predecir el número de comportamientos probables en términos de violencia incontrolable con diferentes valores de severidad; este indicador de acuerdo con Leal-Enríquez podría ampliar su alcance, o conectarse con otros factores, tales como la situación económica, la edad del agresor y de la mujer víctima, a efectos de establecer correlaciones entre ellos; inclusive se afirma que además de ser una herramienta útil en el monitoreo de la violencia, puede ayudar en la formulación de las decisiones de política criminal en relación con la VBG (llamada por el autor violencia de pareja) (Leal-Enríquez, 2018, pág. 1061).

El creador del modelo ha señalado que su motivación fue establecer una relación entre un fenómeno propio de las ciencias sociales y las matemáticas, con el objetivo de establecer parámetros que permitan crear modelos y simuladores de violencia, que puedan ser herramientas útiles para el trabajo de profesionales que trabajen precisamente con ciencias sociales (tales como trabajadores sociales, sociólogos), en el marco por ejemplo de ONG's o asociaciones privadas (LJA.MX, 2022). Así las cosas, una de las principales virtudes de este modelo, es establecer esa conexión, entre dos campos del conocimiento que tradicionalmente se han entendido con opuestos, ya que se permite un acercamiento entre las denominadas ciencias duras (como la matemática), y las ciencias blandas, que abordan asuntos relacionados entre otras, con el comportamiento humano en sociedad.

Así las cosas, una de las principales virtudes de este modelo, es establecer esa conexión, entre dos campos del conocimiento que tradicionalmente se han entendido con opuestos, ya que se permite un acercamiento entre las denominadas ciencias duras (como la matemática), y las ciencias blandas, que abordan asuntos relacionados entre otras con el comportamiento humano en sociedad.

El Gobierno de México ha resaltado que este modelo matemático retoma el concepto de ciclo de la violencia, y que al permitir la simulación de escenarios futuros de violencia, podría ayudar a las mujeres víctimas a percibir con mayor claridad, los escenarios cercanos de violencia a los que estarían expuestas e inclusive el riesgo de muerte (Comisión Nacional para Prevenir y Erradicar la Violencia Contra las Mujeres, 2018).

Aquí, es particularmente llamativo, que esta herramienta sería de gran ayuda, en mujeres que sufren síndrome de indefensión aprendida, o inclusive síndrome de Estocolmo doméstico, los cuales afectan sus capacidades de afrontamiento, y les impiden entre otras, visualizar o aceptar la gravedad del ciclo de violencia en el que se encuentran inmersas.

### C. Otras herramientas

Si bien, la presencia de los algoritmos de predicción de riesgo de violencia basada en género ha empezado a aumentar, se evidencia un mayor crecimiento de otro tipo de instrumentos, que, sin realizar una predicción, si pretenden ayudar en la prevención y lucha contra este fenómeno.

#### a. Certeza de Voz:

El Instituto Andaluz de la Mujer (IAM) y la Empresa Pública de Emergencias Sanitarias, se encuentran trabajando en el desarrollo de una herramienta de detección temprana de violencia contra la mujer, utilizando el patrón vocal (voz, tono, pausas, etc.), para detectar si una mujer es víctima de violencia. Debe anotarse que, debido a la dinámica del ciclo de la violencia, algunas mujeres no logran detectar o aceptar que se encuentran sometidas a violencia basada en género; por ello el proyecto “Certeza de Voz”, pretende que el personal que atiende llamadas recibidas en los Centros de Coordinación de Urgencias y Emergencias Sanitarias de Andalucía cuente con este sistema de procesamiento de información, para detectar estos casos de violencia, y mejorar las posibilidades de atención. Como datos de entrenamiento de esta herramienta, se utilizarán llamadas anteriores de mujeres que no eran conscientes de estar inmersas en una relación de abuso, quienes han prestado su consentimiento para que, de forma anónima, se utilice su precedente (Panda Security, 2021, párr 1-2).

#### b. Project SAFE:

El Project SAFE es un “estudio que aprovecha la inteligencia artificial y la ciencia de datos para mitigar la violencia de género y la inseguridad, durante y después de COVID-19 en Nigeria” (Project SAFE, 2022, párr 1).

Sus creadores reconocen que existen muchas dificultades para las mujeres víctimas de violencia basada en género para denunciar, relacionadas entre otras con la estigmatización y la inequidad en el acceso a canales de denuncia; en ese sentido, recalcan la importancia de utilizar la tecnología para mejorar tales

canales. SAFE es un proyecto desarrollado por la academia de ciencia social de Nigeria y por Towntalk y pretende que, usando herramientas de IA, se logre disminuir el impacto de la VBG, en especial aquella ocurrida durante y con posterioridad a la pandemia derivada del Covid. No puede perderse de vista que la ONU ha denunciado en diversos documentos, el incremento e intensificación que se produjo de la VBG, en el marco de la emergencia sanitaria derivada del Covid, entre otras, por el aumento de presencia de los agresores en el hogar, limitaciones para acudir a la administración de justicia, intensificación de las tensiones en el hogar debido a pérdidas de empleo, disminución de ingresos etc.

En el marco de esa profundización del aislamiento de las víctimas de VBG, SAFE utiliza una aplicación móvil denominada "Area", a través de la cual se proporciona "información verificada sobre riesgos de seguridad cercanos, acceso mejorado a servicios de emergencia (p. ej., refugios) y funciones de seguridad personal", además, se complementa con un servicio denominado "Pulse", el cual es "una herramienta de gestión de casos completamente integrada para personal de primeros auxilios y especialistas de apoyo". (Elusoji, 2021, párr 1,2 ) Resulta interesante que, esta iniciativa combina tanto asistencia dirigida para las víctimas como también herramientas de monitoreo y gestión para los denominados primeros respondientes de los casos, todo ello de la mano de la tecnología.

c. BINDI:

Se trata de un proyecto conducido por UC3M4 Safety, que es un grupo interdisciplinar de la Universidad Carlos II de Madrid (España), y tiene como objetivo la detección, prevención y combate de la violencia contra las mujeres. Este grupo se propuso desarrollar una "un dispositivo portátil no intrusivo, capaz de detectar y alertar automáticamente cuando un usuario está bajo un intenso estado emocional (p. ej., pánico, miedo o estrés) potencialmente causados por una situación de violencia de género, para poder prestar la ayuda correspondiente" (Dziech, Mees & Czyżewsk, 2020, pág. 224)

El dispositivo BINDI se compone de tres elementos, un pendiente y una pulsera (dos dispositivos portables), más una aplicación para teléfonos inteligentes; estos elementos trabajan de la siguiente forma: "el brazalete mide datos fisiológicos, específicamente pulso de volumen sanguíneo (BVP), respuesta galvánica de la piel (GSR) y temperatura de la piel (SKT). El pendiente graba audio y voz principalmente del usuario que lo lleva, pero también del entorno y de personas alrededor" (Dziech et al, 2020, págs. 224-225). Con la información capturada, la IA puede detectar una situación y procede a activar una alarma automática, no obstante, también existe la posibilidad de que la alarma se active manualmente con botones físicos que existen en los dos dispositivos. La alarma automática antes mencionada informa de forma automática a las agencias de control, y las grabaciones de audio son almacenadas en una nube con sistema de encriptación, que garantiza que puedan ser usadas como prueba

en futuros procesos. (Dziech et al, 2020, pág. 225) Con la información capturada, la IA puede detectar una situación y procede a activar una alarma automática; no obstante, también existe la posibilidad de que la alarma se active manualmente con botones físicos que existen en los dos dispositivos. La alarma automática antes mencionada informa de manera automática a las agencias de control, y las grabaciones de audio son almacenadas en una nube con sistema de encriptación, que garantiza que puedan ser usadas como prueba en futuros procesos (Dziech et al, 2020, pág. 225).

Esta herramienta utiliza técnicas de Machine Learning para aprender a detectar emociones alarmantes de peligro en la víctima, iniciando por establecer una medida estándar a los niveles emocionales por medio de las lecturas de los signos vitales que se realizan a través del brazalete. También realiza un análisis con los niveles de voz y sonidos en el entorno, para así validar la voz de la víctima e identificar posibles agresiones verbales o sonidos amenazantes. Con estos dos procesamientos, la IA es capaz de detectar niveles de estrés en la víctima y emitir alarmas en el momento que se presenta la agresión.

3. Elementos (variables) ponderadas por los diferentes sistemas y principales críticas:

Los sistemas mencionados (tanto los que utilizan IA como los que no), usan diversas variables que son ponderadas; SARA por ejemplo como fue descrita, utiliza factores de riesgo tales como historial de delito/agresión, factores psicosociales, entre otros. Dentro del historial delictivo se analizan datos relacionados con la violencia intrafamiliar que haya ejecutado el agresor, sobre el factor de delito se recolecta información relacionada con agresiones de violencia sexual grave u ordenes de alojamiento.

En la misma línea, la herramienta VIOGEN usa las variables relacionadas con violencia de pareja y la asociada con el indicador de elementos psicológicos o psiquiátricos. En lo que se refiere a ODARA igualmente tiene ítems enlazados con actos de violencia, pero esta herramienta actuarial además utiliza ítems vinculados con decisiones judiciales, ejemplo de ello si el perpetrador cuenta con alguna sentencia de tipo correccional, igualmente usa variables como historial de uso de drogas, o historial familiar – biológico (más de un hijo, hijo biológico de la víctima de una pareja anterior) entre otras.

Así, en las diversas herramientas analizadas puede advertirse que los factores o variables ponderadas y valoradas a efectos de establecer niveles de riesgo, combinan los denominados factores estáticos (antecedentes) y factores dinámicos.

Resulta importante señalar que González Álvarez, López Ossorio, & Muñoz Rivas, resaltan que la predicción de riesgo de conducta violenta no es un asunto nuevo; por el contrario,

en los últimos 15 años se han desarrollado nuevas técnicas para valorar y predecir la conducta violenta”, y precisan que el desarrollo de estas herramientas se enfoca de manera prevalente a la prevención, utilizando “tres elementos principales: i) El mejor conocimiento de la naturaleza y procesos que conducen a la violencia ii) La sustitución del concepto jurídico de “peligrosidad” por el de “riesgo de violencia” iii) El desarrollo y aplicación de protocolos e instrumentos de uso profesional para la valoración del riesgo de violencia (2018, pág. 37)

En cuanto al concepto de peligrosidad, tradicionalmente se le ha criticado su falta de precisión y su cercanía a percepciones subjetivas o sesgadas, por lo que el concepto de riesgo, ya la valoración de factores coadyuvantes o disparadores para determinar niveles de riesgos, dentro de las herramientas de valoración de niveles de riesgo, combinan factores relacionados no solo con el eventual victimario, sino también con la víctima, logran aumentar los niveles de precisión, al reconocer que la violencia, es un fenómeno multicausal.

Es así como en el estudio comparativo realizado por Frédérique, Robitaille, Lévesque & Lessard, en el que compararon treinta y cinco (35) herramientas de evaluación de riesgo de homicidio, al abordar lo relativo a los factores utilizados, se señaló que en la escogencia de tales variables se priorizaron “aquellos que podrían ser utilizados por las mujeres en cuestión para evaluar la peligrosidad de su situación respecto al homicidio doméstico sin referirse necesariamente directamente a la violencia doméstica” (2018, pág. 8), (traducción libre).

En la misma línea, el estudio identificó que “los ítems están asociados a las características del autor de la violencia y/o de la víctima, las conductas violentas y coercitivas, y el contexto de la relación conyugal y familiar” (Frédérique, Robitaille, Lévesque, & Lessard, 2018, pág. 8) (traducción libre). Y, como es apenas natural, los autores del mencionado estudio indicaron que estos factores no pueden ser entendidos desde la causalidad, sino como factores detonantes.

Al respecto indicaron que “no hay vínculos causales involucrados aquí ya que la presencia de estos factores de riesgo no indica que ocurrirá un homicidio doméstico. Sin embargo, su presencia aumenta la probabilidad de que pueda ocurrir tal evento” (2018, pág. 8) (traducción libre). En este sentido aquí se acoge lo afirmado en el mencionado estudio, ya que la violencia en general, y la violencia de género en particular, son fenómenos multicausales, en los que no es posible identificar un factor causal único o predominante.

Frédérique, Robitaille, Lévesque & Lessard, clasifican los factores o variables utilizadas en estas herramientas, en cuatro (4) categorías así: i) factores asociados con características del perpetrador de la violencia ii) factores asociados con características de la víctima de violencia iii) contexto de la relación de pareja y contexto familiar iv) Ítems que evalúan conductas violentas o coercitivas.

Es interesante señalar que la mayoría de las herramientas de predicción optan por el modelo combinado de los factores, aunque en muchos de ellos, la ponderación o valor asignado a todas las variables parecer ser igual; esto se advierte principalmente en las herramientas que no usan inteligencia artificial. En este tipo de herramientas “manuales”, o con participación exclusivamente humana, se opta por calificar las diversas variables con puntajes dentro de un rango, para luego totalizar el puntaje, que a su vez llevará al establecimiento de un nivel de riesgo; aquí se echa de menos entonces, el diferente peso que puede atribuirse a cada factor.

De lo anterior, también podría desprenderse la dificultad de las herramientas de predicción que utilicen IA, para determinar el contenido y/o peso (incidencia) de cada factor en el puntaje final, lo que además puede resultar incierto cuando quiera que la IA actúe con cierta autonomía como consecuencia del machine learning. A lo anterior, se suman las dificultades que ya se han señalado en artículos previos (Roa Avella, Sanabria - Moyano, & Dinas - Hurtado, 2022) relacionadas con el concepto de black box, y la dificultad de acceder a los procesos internos que ocurren al interior del algoritmo, previos a la obtención del resultado de predicción de nivel de riesgo, lo que genera preocupaciones al no poder establecer la forma exacta en la que funciona la herramienta y de esta forma identificar si las variables que usa tienen o no algún tipo de sesgo.

Ahora bien, en particular respecto de la herramienta ODARA, se han precisado las siguientes debilidades:

- ODARA no incluye todos los factores validados con evidencia para predecir alto riesgo o letalidad
- A pesar de que un individuo obtenga un puntaje bajo, ello no excluye riesgo de tendencias suicidas o de agresión sexual
- Se advierte que una puntuación baja puede derivarse de las dificultades en el acceso a la información relacionada con los antecedentes penales del posible agresor
- También se resalta que el recaudo de información puede verse afectado por razones culturales o de barreras de lenguaje (Maine Guidelines & General Scoring Criteria - ODARA)

Finalmente, en este tema, es necesario referir que en materia de las variables, factores o ítems, el modelo matemático creado por el investigador mexicano Erik Leal-Enríquez, ha introducido la novedad del criterio de auto control (o pérdida de control) del perpetrador de la violencia, el cual no se encuentra identificado en ninguna de las herramientas analizadas en este escrito, y en general en la mayoría de las identificadas por la doctrina, por lo que su uso permitirá a futuro establecer las bondades o dificultades de su incorporación.

Podemos concluir este apartado señalando que, cualquiera que sea el instrumento escogido, los factores que éste utilice para la predicción, o si este

es exclusivamente humano, o utiliza IA; ello no elimina la presencia de ciertas limitaciones, ya que hasta la fecha, no es posible llegar a una predicción exacta o absolutamente confiable; además en algunas herramientas al no poder saber la forma exacta en que funcionan (ya que se limita el acceso a la misma) se deja en evidencia lo que se ha mencionado antes relacionado con la black box. Sin embargo, la escogencia de los factores o variables, al igual que las formas de ponderación, van a afectar el grado de precisión de la evaluación de riesgo.

#### 4. Ventajas y Desventajas del uso de herramientas de predicción de niveles de riesgo en materia de VBG

La IA ha traído consigo grandes avances en la toma de decisiones en el sector legal, en ese sentido los algoritmos “[...] son herramientas que aportan celeridad a la justicia y precisión en la toma de decisiones [...]” (Roa Avella et al, 2022, pág. 277), de tal forma que, con el uso de las herramientas de predicción de niveles de riesgo en materia de VBG que usan IA y aquellas que no, se destacan grandes ventajas, relacionadas principalmente con la posibilidad de mejorar la investigación, judicialización y sanción de la VBG, a través del establecimiento de niveles de riesgo, que a su vez implica probabilidad de ocurrencia de la conducta violenta. Los usos y ventajas se alinean principalmente con prevención, planes de seguridad y medidas de protección individualizadas.

Se ha precisado que el desarrollo e implementación de sistemas computarizados de apoyo a las decisiones (DSS) y las herramientas de evaluación de riesgos que se basan en datos estandarizados (dentro y/o entre agencias), ayudan a comprender el riesgo de reincidencia de DV (domestic violence) para subgrupos dentro de la población” (McNamara, Graham, Broad, & Soon Ong, 2018, págs. 3,4). Respecto de las experiencias con los algoritmos predictivos de niveles de riesgo, de acuerdo con McNamara, Graham y Broad, pueden resaltarse como ventajas, el aumento en los niveles de precisión en las decisiones judiciales, mayor transparencia en las mismas, y mejor acceso a datos de entrada.

Si quisieran señalarse ventajas particulares de las herramientas antes señaladas, se puede afirmar que mientras que SARA analiza y recolecta información del posible agresor (historial delictivo, historial de violencia en pareja, historial psicosocial, entre otros) con el fin de brindar una valoración final, línea el algoritmo VIOGÉN además de predecir el riesgo, se enfoca en proteger a las víctimas, a lo que se le suma la ventaja de la adición reciente del software SAS IBERIA, con el que se acelera el análisis de la información recolectada del algoritmo VIOGÉN, lo que conlleva entonces a atender de forma más rápida algunos casos donde la víctima se encuentra en riesgo.

Lo anteriormente mencionado deja en evidencia entonces, que algunas de las utilidades de estos instrumentos apuntan a la velocidad y facilidad en la



recolección, análisis y valoración de grandes cantidades de datos, lo cual no sería posible con mera participación humana debido a nuestras limitaciones cerebrales; por otra parte, esa información es la materia prima de diversas actividades que tienen como objetivo proteger a la víctima y a su vez salvaguardar todos aquellos derechos que podrían estar siendo vulnerados a causa de estas conductas de VBG .

Por otra parte, sobre las desventajas de estas herramientas, diversos estudios han hecho énfasis en las debilidades del uso de la IA en la evaluación de riesgo, en particular señalando el secretismo, la complejidad innecesaria y el reforzamiento de sesgos existentes; lo anterior, en relación con la denominada black box, que como ya se ha clarificado, representa la imposibilidad de acceder a los procesos previos al establecimiento del nivel de riesgo, a lo que se añade la complejidad antes mencionada, la cual se relaciona con el lenguaje propio del algoritmo, la magnitud de los datos utilizados y procesados por aquel, que son imposibles de entender por parte de la mente humana.

No puede pasarse por alto, la posibilidad de presencia de sesgos, como se señaló respecto la herramienta actuarial ODARA, la cual ha sido tildada de estar sesgada respecto de las comunidades indígenas, o minorías, lo que afectaría la confiabilidad y precisión de los resultados, los cuales, de paso, vulnerarían diversos derechos, y en particular, atentarían contra lo establecido en el artículo 7 de la Declaración Universal de DDHH, en relación con el derecho a no discriminación, presunción de inocencia e inclusive derecho a la igualdad. Lo anterior devela una preocupación, en torno al conflicto entre la ayuda que estas herramientas ofrecen, en la lucha por la eliminación de la VBG, versus los posibles riesgos de vulneración de otros derechos y garantías establecidas en los instrumentos internacionales.

Teniendo en cuenta las críticas mencionadas sobre estas herramientas, han surgido propuestas tales como el diseño ético de los algoritmos. En ese sentido existen organizaciones cuyo propósito es luchar en contra de las injusticias algorítmicas. Puede resaltarse, el esfuerzo de la activista Joy Buolamwini, quien ha creado la Liga de la Justicia algorítmica, y ha señalado que en el mundo actual la IA gobierna el acceso a la información, las oportunidades y la libertad, puntualizando que los sistemas inteligentes pueden perpetuar el racismo, la discapacidad, y en general muchas formas de discriminación. Por ello propone la denominada “codificación inclusiva”, para moderar los efectos de la subrepresentación de mujeres y personas de color en el sector de la tecnología (Buolamwini, párr. 1, como se cita en Roa Avella, Sanabria - Moyano, & Dinas - Hurtado, 2022, pág. 303).

Una recomendación que parece prudente en relación con las herramientas predictivas que usan IA, es que la información con la que son alimentadas pueda ser pública, ya que los proveedores de algunos sistemas son empresas privadas que limitan los datos que comparten de forma libre; así, si la información fuese

pública se combatiría la black box, lo cual además estaría contribuyendo a un diseño transparente y ético de estas herramientas algorítmicas. En la misma línea es pertinente insistir en que dichas herramientas sean construidas garantizando que están libres de sesgos hacia comunidades específicas.

No obstante, las diversas críticas, es indudable que la presencia de herramientas que utilicen IA abre un nuevo espacio en la prevención, judicialización y sanción de la VBG, adicionalmente, las herramientas diseñadas para facilitar el acceso de las víctimas a canales de ayuda, botones silenciosos, bots y demás, representan una enorme oportunidad para combatir estos escenarios de violencia.

#### 6. Variables sugeridas para la construcción de herramientas de IA que busquen predecir niveles de riesgo de VBG y/o feminicidio

En el campo de la predicción de riesgo de violencia es importante partir de que aquella es un fenómeno pluricausal, y que “es entendida como el resultado de una compleja interacción de factores, tanto ambientales y sociales (maltrato infantil, abuso de sustancias, experiencia previa, cultura, etc.) como biológicos” (Moya-Albiol, Sariñana-González, Vitoria-Estruch, & Romero-Martínez, 2017, pág. 16) En la categoría de biológicos, los anteriores autores determinan que se incluyen “los factores neuroanatómicos y neuroquímicos, los hormonales, los genéticos y los moleculares”; respecto de estos últimos debe tenerse en cuenta que pueden interactuar con la violencia, ya sea como predisponentes, como detonantes, o también pueden aparecer o modificarse con posterioridad al episodio violento (Moya-Albiol et al, 2017, pág. 16).

Si bien parece existir una cierta uniformidad en los factores o grupos de factores ponderados por las herramientas de predicción, las recomendaciones al respecto deben alinearse con la necesidad de proteger los derechos de las víctimas; en ese sentido, las variables relacionadas con ellas deben alejarse de la culpabilización de la víctima, y de la revictimización. En este tema, los sistemas unificados de información de las diversas agencias que tienen contacto con mujeres víctimas de VBG son de vital importancia, al igual que la presencia de funcionarios con visión de género, que tengan la capacidad de identificar las dificultades que para estas mujeres puede suponer el compartir información, o establecer su verdadera situación de riesgo.

En el recorrido que se ha hecho por los principales elementos de predicción de niveles de riesgo, las herramientas tradicionales que no utilizaban IA, se centraban en grupos de factores bastante similares, en general dentro de la categoría de factores estáticos (tales como los antecedentes, nivel de escolaridad, vinculación laboral, entre otros), en tanto que la entrada de herramientas o procesos que involucran I.A, abrió el paso a otros factores considerados más dinámicos, tales como aquellos relacionados con estados anímicos, control de impulsos, agresividad, etc. Así, se ha abierto el paso para

que inclusive hacia futuro puedan ponderarse factores neuro-conductuales o genéticos.

En el escenario de la predicción de violencia, la neurocriminología ha empezado a abrirse paso como una novedosa herramienta. Aquella pretende aplicar la metodología y técnicas de las neurociencias, a efectos no sólo de comprender el fenómeno violento, sino de indagar cómo predecirlo y prevenirlo (2017, pág. 17). Se trata entonces de un campo interesante en el que con “carácter prospectivo se pretende analizar si factores biológicos concretos podrían ser predictores de la violencia, la delincuencia y la criminalidad” (Moya-Albiol et al, 2017, pág. 17).

Si bien ya se ha iniciado este camino, el futuro estaría relacionado con la predicción de la reincidencia del comportamiento violento, mezclando en la ponderación factores disparadores y factores protectores (crimino-repelentes) frente a la ocurrencia del fenómeno violento. Con esto, una formulación sencilla al respecto sería que

a mayor número de factores relacionados con la violencia o factores de predisposición o vulnerabilidad (alteraciones en circuitos neurales, bajo cortisol y serotonina, alta testosterona, etc.), mayor probabilidad de desarrollar la misma, mientras que el desarrollo de factores protectores como el trabajo en empatía, las técnicas de control de impulsos, la intervención sobre los déficits cognitivos, etc., disminuiría la incidencia de actos violentos (Moya-Albiol et al, 2017, pág. 18).

A pesar de lo atractiva que resulta la incorporación de estos novedosos factores y ciencias, al tratarse de factores médicos, surgiría la dificultad de vincularlos a la ponderación del algoritmo predictivo; no por la parametrización (que puede hacerse de la mano de los estándares médicos), sino por tratarse de datos sensibles, privados o semiprivados (según se trate de la legislación del país), cuyo acceso demandaría una orden judicial. Sin embargo, podría evaluarse que, cuando quiera que existan antecedentes (condenas) por delitos violentos, las autoridades tengan la posibilidad de ordenar la inserción de estos indicadores en los formularios de evaluación de riesgo. Este asunto, sin duda podría generar resistencias ya que, desde el lado del presunto agresor, puede verse como una injerencia indebida, constitutiva además de una violación al derecho a la intimidad, que debería zanjarse con un ejercicio de ponderación (al mejor estilo de Robert Alexy), en el que el riesgo para la salud y/o la vida de la mujer saldrían victoriosos.

Finalmente, el futuro de las herramientas predictivas abre un espacio importante para luchar en contra de la parcelación del conocimiento, para plantear escenarios de interdiscipliniedad que acerquen ciencias tan disímiles como la matemática, la estadística, la ingeniería, de un lado; y del otro algunas de las denominadas ciencias sociales, tales como la sociología o la antropología. Sin embargo, ello revive también el debate acerca de la dificultad de “predecir”

el comportamiento humano, en tanto se aleja de procesos causales, en atención a la presencia entre otros, de la racionalidad, libre albedrío e inteligencia humana.

### **Conclusiones**

La IA ha tenido grandes avances y es así como en la implementación de técnicas como el ML, se han desarrollado algoritmos de predicción de riesgo de violencia que ponderan factores diversos con miras a establecer la probabilidad de ocurrencia de VBG y/o feminicidio, lo que ha permitido entonces que este riesgo disminuya y a su vez se puedan proteger los derechos que puedan llegar a vulnerarse con estas acciones ejecutadas por parte del perpetrador.

También existen herramientas tradicionales, que, al no utilizar IA, incorporan un juicio/evaluación humana, de tipo clínico normalmente. Estas herramientas siguen siendo un recurso útil, relativamente sencillo, que ayuda en la determinación de niveles de riesgo, principalmente de reincidencia; sin embargo, la permanente crítica en torno a fiabilidad, presencia de sesgos y posibilidad de falsos positivos y negativos, sigue siendo su talón de Aquiles.

Por otra parte, es esperanzador ver el desarrollo de diversas herramientas que se han mencionado en precedencia tales como Certeza de voz y BINDI, que representan el esfuerzo de muchos sectores de la sociedad, por establecer canales de ayuda rápida, sencilla y eficaz, para que las víctimas de VBG puedan recibir atención oportuna por parte de las autoridades, y de esta manera disminuir así sea parcialmente, los diversos tipos de violencia contra las mujeres y los feminicidios.

Finalmente, el futuro de las herramientas predictivas de riesgo de conducta violenta se enfoca hacia las neurociencias, la neurocriminología, y en general los eventuales desarrollos que involucren análisis de aspectos que parecían enterrados desde la época del muy criticado Lombroso, pero que hacen su entrada triunfal, de la mano de tecnologías cada vez más precisas, que sin embargo regresan sobre la gran pregunta de qué es lo que genera la conducta delictiva.

Sea cual sea la herramienta desarrollada, los factores escogidos para la ponderación, los datos de entrenamiento y los parámetros establecidos por el desarrollador, lo que no puede perderse de vista es la necesidad de incorporar a estos procesos la ética algorítmica, la auditoría constante y el análisis de la presencia de sesgos, con miras a mantener una tecnología al servicio del hombre, respetuosa de los DDHH, y en general de la dignidad humana.

## Bibliografía

- González-Prieto, Á., Brú, A., Nuño, J. C. & González-Álvarez, J. L. (2021). Machine learning for risk assessment in gender-based crime. *ArXiv*, *abs/2106.11847*. <https://doi.org/10.48550/arXiv.2106.11847>
- Elusoji, S. (14 de mayo, 2021). Project SAFE, Partners To Bolster Support Systems For GBV Survivors. *Channels*. <https://www.channelstv.com/2021/05/14/project-safe-to-hold-convening-to-bolster-support-systems-for-gbv-survivors/>
- Cheishvili, A. (25 de agosto, 2021). The AI Revolution Is Happening Now. *Forbes*. <https://www.forbes.com/sites/forbestechcouncil/2021/08/25/the-ai-revolution-is-happening-now/?sh=42d2e75c28c8>
- Liu, Z. & Chen, H. (2017). *A predictive performance comparison of machine learning models for judicial cases*. 2017 IEEE Symposium Series on Computational Intelligence (SSCI). <http://dx.doi.org/10.1109/SSCI.2017.8285436>
- Ferguson, C. & McLachlan, F. (2020) *Predicting and assessing lethal risk in domestic and family violence situations in Australia*. QUT Centre for Justice. <https://eprints.qut.edu.au/216824/>
- Flores, S. (14 de enero, 2021). Inteligencia artificial en la valoración del riesgo para las víctimas de violencia de género. *Observatorio Violencia*. <https://observatorioviolencia.org/inteligencia-artificial-en-la-valoracion-del-riesgo-para-las-victimas-de-violencia-de-genero/>
- Frédérique, L., Robitaille, C., Lévesque, S., & Lessard, G. (2018). *Domestic Homicide Risk Assessment: An Inventory of Existing Tools*. Équipe de recherche en violence conjugale. [https://www.raiv.ulaval.ca/sites/raiv.ulaval.ca/files/publications/fichiers/fiche\\_synthese\\_13\\_eng\\_finale\\_web.pdf](https://www.raiv.ulaval.ca/sites/raiv.ulaval.ca/files/publications/fichiers/fiche_synthese_13_eng_finale_web.pdf)
- Comisión Nacional para Prevenir y Erradicar la Violencia Contra las Mujeres. (2 de febrero, 2018). ¿Es posible medir matemáticamente la violencia contra las mujeres? *Gobierno de México*. <https://www.gob.mx/conavim/articulos/es-posible-medir-matematicamente-la-violencia-contra-las-mujeres?idiom=es>
- González-Álvarez, J. L., López-Ossorio, J. J., & Muñoz-Rivas, M. (2018). *La valoración policial del riesgo de violencia contra la mujer pareja en España–Sistema VioGén*. Ministerio del Interior, Gobierno de España. [https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/seguridad-ciudadana/La\\_valoracion\\_policial\\_riesgo\\_violencia\\_contra\\_mujer\\_pareja\\_1261808\\_87.pdf](https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/seguridad-ciudadana/La_valoracion_policial_riesgo_violencia_contra_mujer_pareja_1261808_87.pdf)
- Granados, Ó. (18 de febrero, 2021). La Inteligencia Artificial Ayuda a Prevenir la Violencia de Género. *EL PAÍS*. <https://elpais.com/tecnologia/haz-cosas->

[extraordinarias/2021-02-18/la-inteligencia-artificial-ayuda-a-prevenir-la-violencia-de-genero.html](https://doi.org/10.1177/0008125619864925)

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), pp. 5–14. <https://doi.org/10.1177/0008125619864925>

Hegel, J., Pelletier, K. D., & Olver, M. E. (2022). Predictive Properties of the Ontario Domestic Assault Risk Assessment (ODARA) in a Northern Canadian Prairie Sample. *Criminal Justice and Behavior*, 49(3), pp. 411–431. <https://doi.org/10.1177/00938548211033631>

IBM Cloud Education. (3 de junio, 2020). Inteligencia artificial (IA). IBM. <https://www.ibm.com/co-es/cloud/learn/what-is-artificial-intelligence>

Dziech, A., Mees, W. & Czyżewsk, A. (2020). *Multimedia Communications, Services and Security*. [10th International Conference, MCSS 2020, Kraków, Poland, October 8-9, 2020, Proceedings]. Springer.

Leal-Enríquez, E. (2018). Mathematical modeling of intimate partner violence: Simulations of loss of control scenarios. *Journal of Computational and Applied Mathematics*, 330, pp. 1052-1062. <https://doi.org/10.1016/j.cam.2017.07.002>

LJA. (15 de febrero, 2016). Desarrollan Modelo Matemático Para Predecir Futuro De Violencia Hacia Las Mujeres. LJA.MX <https://www.lja.mx/2016/02/desarrollan-modelo-matematico-para-predecir-futuro-de-violencia-hacia-las-mujeres/>

McNamara, D., Graham, T., Broad, E., & Soon Ong, C. (2018). Trade-offs in Algorithmic Risk Assessment: an Australian Domestic Violence Case Study. En A. Daly, M. Mann, & S. K. Devitt (Eds.) *Good data (Theory on Demand, 29)*, pp. 96-116. Institute of Network Cultures, The Netherlands.

Lee, M. K., Jain, A., Cha, H. J., Ojha, S. & Kusbit, D. (2019). Procedural Justice in Algorithmic Fairness: Leveraging Transparency and Outcome Control for Fair Algorithmic Mediation. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 182, pp. 1-26. <https://doi.org/10.1145/3359284>

Ministerio del Interior. (02 de febrero, 2022). Sistema VioGén. <https://www.interior.gob.es/opencms/ca/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen/>

Moya-Albiol, L., Sariñana-González, P., Vitoria-Estruch, S., & Romero-Martínez, Á. (2017). La neurocriminología como disciplina aplicada emergente. *Vox Juris*, 33(1), pp. 15-20. <https://dialnet.unirioja.es/servlet/articulo?codigo=6058766>

Hilton, N. Z., & Radatz, D. L. (2021). Criminogenic needs and intimate partner violence: Association with recidivism and implications for treatment. *Psychological services*, 18(4), pp. 566–573. <https://doi.org/10.1037/ser0000450>

Organización de los Estados Americanos, OEA. (9 de junio de 1994). *Convención Interamericana Para Prevenir, Sancionar Y Erradicar La Violencia Contra La Mujer*

"Convención De Belém Do Pará".

<https://www.oas.org/es/cidh/mandato/Basicos/13.CONVENCION.BELEN%20DO%20PARA.pdf>

ONU Mujeres. (2022). *Hechos y cifras: Poner fin a la violencia contra las mujeres*. <https://www.unwomen.org/es/what-we-do/ending-violence-against-women/facts-and-figures>

Organización Mundial de la Salud & Organización Panamericana de la Salud. (2013). *Comprender y abordar la violencia contra las mujeres: violencia sexual*. Organización Mundial de la Salud. <https://apps.who.int/iris/handle/10665/98821>

Panda Security. (12 de abril, 2021). La inteligencia artificial servicio de la lucha contra la violencia de género. *Panda media center*. <https://www.pandasecurity.com/es/mediacenter/mobile-news/inteligencia-artificial-violencia-genero/>

Pueyo, A., López, S., & Álvarez, E. (2008). Valoración del riesgo de violencia contra la pareja por medio de la SARA. *Papeles del psicólogo*, 29(1), pp. 107-122. <https://dialnet.unirioja.es/servlet/articulo?codigo=2536843>

Roa Avella, M. D., Sanabria-Moyano, J., & Dinas-Hurtado, K. (2022). Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos. *Revista Brasileira de Direito Processual*, 8(1), pp. 275-311. <https://doi.org/10.22197/rbdpp.v8i1.615>

SAS. (15 de diciembre, 2020). *La Secretaría de Estado de Seguridad y SAS se unen en la lucha contra la violencia de género con analítica avanzada e inteligencia artificial*. [https://www.sas.com/es\\_es/news/press-releases/locales/2020/viogen-secretaria-estado-seguridad-y-sas-unidos-lucha-contra-violencia-genero-analitica-avanzada-ia.html](https://www.sas.com/es_es/news/press-releases/locales/2020/viogen-secretaria-estado-seguridad-y-sas-unidos-lucha-contra-violencia-genero-analitica-avanzada-ia.html)

*The Maine Coalition to end domestic violence y Violence Intervention Partnership of Maine Cumberland County*. (2014). Maine Guidelines & General Scoring Criteria ODARA.

Ulmer, J. C. (2015). The Ontario Domestic Assault Risk Assessment (ODARA): A Validation and Comparison Study for an Oregonian Law Enforcement Agency. *Doctor of Psychology (PsyD)*, paper 177. <https://digitalcommons.georgefox.edu/psyd/177/>

UN Women. (s.f.). *Global Database on Violence against Women: Colombia*. <https://evaw-global-database.unwomen.org/en/countries/americas/colombia>

Yablon, C., & Landsman-Roos, N. (2013). Predictive Coding: Emergin Questions and Concerns. *South Carolina Law Review*, 64(3), pp. 633-680. <https://scholarcommons.sc.edu/sclr/vol64/iss3/6>