



The relevance of the data collection process in the VioGén system from a feminist perspective

OÑATI SOCIO-LEGAL SERIES VOLUME 14 ISSUE 6 (2024), 1555–1573: THE INFLUENCE OF NEW TECHNOLOGIES ON LAW

DOI LINK: [HTTPS://DOI.ORG/10.35295/OSLS.IISL.1922](https://doi.org/10.35295/OSLS.IISL.1922)

RECEIVED 20 NOVEMBER 2023, ACCEPTED 2 JULY 2024, FIRST-ONLINE PUBLISHED 22 JULY 2024, VERSION OF RECORD PUBLISHED 1 DECEMBER 2024

JAVIERA FARIÁS PEREIRA*

Abstract

The VioGén system is a predictive policing and algorithm tool, which has been implemented in Spain since 2007. Its main objective is to assess the risk of a woman of being victim of intimate partner violence and proposing measures for her protection accordingly. This work will focus on one of the most essential parts of the tool's functioning: the data collection process. I intend to illustrate how this process of data collection can be influenced by the bias and stereotypes that have been constructed within criminal law concerning gender violence victims and the way this affects, in the end, the system's functioning.

Key words

Algorithms; predictive policing; criminal justice; gender; discrimination

Resumen

El sistema VioGén es una herramienta algorítmica y policial predictiva que lleva implantada en España desde 2007. Su principal objetivo es evaluar el riesgo de que una mujer sea víctima de violencia de pareja y proponer medidas para su protección como consecuencia de ello. Este artículo se centrará en una de las partes más esenciales del funcionamiento de la herramienta: el proceso de recopilación de datos. Pretendo ilustrar cómo este proceso de recolección de datos puede verse influido por los sesgos y estereotipos que se han construido dentro del derecho penal sobre las víctimas de violencia de género, y la forma en que esto afecta, en última instancia, al funcionamiento del sistema.

* Javiera Farías Pereira is Predoctoral researcher at the Department of Philosophy of Law of the University of Barcelona. Email: jfariape29@doct.ub.edu

Palabras clave

Algoritmos; vigilancia predictiva; justicia penal; género; discriminación

Table of contents

1. Introduction	1558
2. Background: Predictive policing systems and the incorporations of algorithms ...	1559
3. Women and uses of technology: The threat of bias and discrimination	1560
4. Women and penal system: The suspicious of feminism.....	1563
5. Algorithmic tools implemented in Spain: The case of VioGén.....	1564
5.1. The Viogén system in perspective.....	1565
5.2. How VioGén works: Éticas Foundation Audit.....	1567
5.3. Data. Halfway between the record and the result	1568
6. Final reflections.....	1570
References.....	1571

1. Introduction

The incorporation of decision-making tools based on algorithms in public administration is, nowadays, common and extensive. In Spain, these systems are used in areas such as employment, prisons, police, education, social services, and fraud detection, among others (Jiménez Arandia 2023, p. 22).

Risk assessment tools stand out amongst the various options available. Great trust has been placed in them, especially in the criminal justice field, due to their ability to predict future behaviour and standardise the decision-making process. These tools are based on an actuarial logic that seeks to create individualised decisions based on a massive database of aggregated information (Harcourt 2007, p. 10). One of the main applications of these tools is in predictive policing systems, which seek to assess the risk of different criminal behaviours. They have been implemented in several countries for an extended time to support the work of the police.

In the last few years, technological development has made it possible to incorporate algorithmic systems to improve the processing of data, which has increased the efficacy of the tools and, at the same time, has contributed to a growing optimism related to their utility and predictive capacity. In Spain, the Viogén system is a good example of an algorithmic predictive policing tool. It was created to assess the risk of women who complain of Intimate Partner Violence to suffer these crimes again and has been implemented in most of the autonomous communities of Spain (except for Catalonia and Basque Country) since 2007.

However, when addressing the criminal system's functioning and women's participation, we cannot forget that this interaction is far from neutral. Still, it is constructed over ideas that put women in specific places and roles. In this sense, the theoretical development made by gender studies and feminist criminology has shown that the law reflects the priorities of a patriarchal order (Naffine 1990, Smart 1994, Pitch 2003). That is why the position of women has always been complicated but, in most cases, subordinated to masculine interests (Naffine 1990, p. 13). In this framework, the law performs the construction of categories to operate: creates a woman's concept, as opposed to a man's, and considering that conceptualisation, the different types of women, such as the criminal woman, the prostitute, the mother, among others (Smart 1994, p. 180). That is how the idea of a perfect victim has been created, according to which the law elaborates their institutions and guides their customs (Pitch 2009, p. 120).

To the old problems related to the place of women in criminal law we must add the new problems of predictive policing system and the use of algorithms on it. With the growing technological evolution in this area, various voices have emerged that warn us of the dangers that the use of algorithms can generate, including concerns about the transparency of the system, the traceability of its operation and the difficulty of attributing responsibility for its results (Mittelstadt *et al.* 2016). There is also a risk of bias and discriminatory results, a risk that can be generated in relation to the data used to operate the system (what data is used, how it is collected), how this data is processed (based on what criteria, what relevance some factors have over others) and in terms of the final result (Coeckelbergh 2021, pp. 107–111). For feminism, meanwhile, the debate on the use of technology has been based on two opposing ideas: the possibility that it

can become an emancipatory tool or that, like many others, it only reinforces positions of subjugation of women (Wajcman 2006).

In this context, the use of predictive policing tools has threatened the right to equality and non-discrimination of women, mainly because of the possibility that the treatment they receive, and the decisions made about them are guided by biases and prejudices – a long-standing problem in the criminal justice system that is renewed by the multiplying effect of the use of algorithms.

This work focuses, firstly, on analysing how data collection process for the functioning of algorithm predictive policing tools can be influenced by the bias and stereotypes that have been constructed in criminal law concerning gender violence victims. Secondly, it seeks to illustrate how this could lead to discriminatory responses. In achieving these goals, the VioGén system will be used as an example to show how these issues operate.

First, I will describe the theoretical source of this kind of tool and how it has been reinforced by the incorporation of algorithms and the new challenges this generates. Secondly, I will explain the theoretical framework developed for the feminist jurisprudence and feminist criminology about the position of women in criminal law, followed by the old and new debates about the use of technology related to gender issues. Then, I will draw a brief characterisation of the VioGén system and the evaluation of its functioning by the Éticas Foundation in 2021. The report of this foundation will be the starting point of the analysis of the relation between the problems with the data collection in this tool and the conceptualisation of the intimate partner violence victims by criminal law.

2. Background: Predictive policing systems and the incorporations of algorithms

Among the risk prediction tools used in the criminal justice system, predictive policing has an essential notoriety and relevance. The concept of ‘predictive policing’ brings together a range of analytical tools that anticipate where and how crimes will be committed (Bennett Moses and Chan 2018, p. 806). The main objectives of these tools are to identify potential perpetrators of a crime, potential victims, and to predict where and when a crime is most likely to be committed again (Hardyns and Rummens 2018, p. 203, González-Álvarez *et al.* 2020, p. 27).

Predictive policing tools are often used to improve the prevention efforts of police forces and to optimise the use of their resources (González-Álvarez *et al.* 2020, p. 27), and their instruments range from the use of basic statistics or simple mechanisms for organising information to advanced technology that allows the processing and visualisation of a large amount of data (Bennett Moses and Chan 2018, p. 808). Predictive policing systems have evolved as technology has advanced, so today, many of them have incorporated algorithmic systems for data analysis, allowing them to process more information and increase their scope.

Simply put, algorithms are instructions given to a system to perform a specific task (López Baroni 2019, p. 10). Considering this starting point, algorithms can have different degrees of complexity. In addition to technical aspects, algorithms have political and social dimensions. They are programmed by humans resulting in their design

incorporating personal worldviews. These technologies incorporate and promote the values that are part of their design, either explicitly or unconsciously (Surden 2013, p. 1). It is therefore argued that they are culturally, historically and institutionally situated (Kitchin 2017, pp. 17–18).

Therefore, the use of algorithms in predictive policing systems brings with it new challenges. These challenges manifest themselves less in qualitative terms since the main problems are related to predictive policing, such as lack of transparency, bias and discrimination. However, with algorithms these problems are exacerbated as the amount of data an algorithm can process and the capacity to respond to it is such that using these technologies amplifies all the risks that predictive policing entails.

3. Women and uses of technology: The threat of bias and discrimination

Feminist studies on technology have focused on how technology influences the construction of gender relations and the place of women in this scheme. Feminists initially adopted a pessimistic position on technology. Radical feminists, ecofeminists and cultural feminists considered technology to be eminently patriarchal, as it was constructed by men and built on male values, and its use was designed to maintain and reinforce women's subordinate position (Wajcman 2006, p. 33). Feminist socialists, on the other hand, focused the problem on men's monopoly on technology, which became an important source of power. Focusing on the division of labour, they believed that this power derived from the lack of women's participation in the technology industry, both in the creation of technology and in jobs in this sector, with women being relegated to domestic or less technical work (Wajcman 2006, pp. 44–46).

From the 1990s onwards, the pessimistic view of technology began to give way to a more optimistic view of the position of women in this context, opening the possibility of making emancipatory use of it (Wajcman 2006, p. 53). During this period, the idea of the social construction of technology began to be studied, understanding it as a product with a strong social influence, both in its use and in its technical content (Wajcman 2006, p. 55). In this sense, “[s]ocial studies of technology emphasise that what favours one technology over another is not necessarily technical efficiency, but rather the set of socio-technical circumstances and the interplay of institutional interests” (Wajcman 2006, p. 58).

An example of this phenomenon could be seen on predictive policing systems, since its use and the implementation of the resulting measures will depend on the level of police confidence in the system and, therefore, on how willing they are to orient their work according to those predictions (Bennett Moses and Chan 2018, p. 814). If the system is opaque, the agents who use it cannot understand how it works, which leads to a lack or an excess of trust in it (González-Álvarez *et al.* 2020, p. 28). In the first case, the user could disregard the information provided and choose to rely on his or her judgement to make the decision; in the second case, the result would be the complete opposite, i.e., the person would disregard his or her judgement and delegate decision-making to the machine (González-Álvarez *et al.* 2020, p. 28).

In this context, feminist perspectives gave content to the analysis of the sociology of technology, opening the debate about how this social influence determined that technology responded to a particular gender order (Wajcman 2006, 67). Through this

exercise, they uncovered the fact that the lack of questioning of gender relations was because when masculinity was dominant, it made the gender question invisible, which appeared only when women entered the equation (Wajcman 2006, p. 73).

In this sense, it is argued that the conception of technology as a neutral entity responds to the fact that when it is controlled by dominant groups, who are not afflicted by any form of oppression, their ideas and ways of seeing the world become dominant. In this context, race, class and gender are not relevant, as these are matters of otherness (D'Ignazio and Klein 2020, pp. 59–60).

In relation to algorithmic systems this situation is similar. Their creation is controlled by people in hegemonic positions who pass on their worldview to them and for whom it is very difficult to visualise the oppressions that affect other, less privileged people because they have never experienced them (D'Ignazio and Klein 2020, p. 28). This phenomenon has been called “risk of privilege”, which also makes it difficult for them to consider these oppressions in their models, either to prevent the harms they may generate in this sense or to anticipate solutions to them (D'Ignazio and Klein 2020, p. 31).

This explains part of the biases that systems incorporate. Another source of bias is the data used to operate the systems, which are not always representative of the reality they are intended to analyse. In some cases, the lack of representation of certain groups prevents solutions adapted to their situation from being reached. This is also the case with predictive policing systems, where it is assumed that the data they use accurately reflect reality, but the true is that not all crimes are equally represented since there is a black figure of those that are not reported (González-Álvarez *et al.* 2020, pp. 28–29). This means that there is always a part of reality that is not reflected in historical data, or that cannot be directly observed by the police (Bennett Moses and Chan 2018, p. 809).

Another possibility is that there is an over-representation of certain situations or groups in the data. This over-representation of data often results from hyper-surveillance of minority groups by large institutions (D'Ignazio and Klein 2020, p. 39). This is often the case with data on crime hotspots used in predictive policing systems: the systems use historical police records, which already define certain neighbourhoods and profiles as more prone to crime, therefore, police activity will be focused on them, increasing the likelihood that crimes committed in these places or by these people will be discovered (Strikwerda 2021, pp. 429–430). This generates a “feedback loop” in the system, which shapes a new reality based on the biased data it was fed, confirming and reinforcing the initial bias (O'Neil 2017, p. 21).

At the same time, the recording of the information collected will not necessarily be accurate or based on the same criteria: many discretionary decisions are made at this point, such as what is defined as a crime, how to interpret the severity or relevance of the information that merits recording, among others, which becomes more apparent when different sources of information are combined (Bennett Moses and Chan 2018, p. 810).

The bias in the data is directly related with the discriminatory outcomes. Although the two concepts are often treated together, they are two distinct situations occurring at different points in the decision-making process. Bias refers to the partial views within the decision-making process. At the same time, discrimination manifests itself in the

outcome of the decision if it disproportionately affects certain groups of people (Mittelstadt *et al.* 2016, p. 8). In this sense, bias can be an antecedent of a discriminatory decision, hence their close linkage and the joint treatment given to them in most cases.

Discrimination can be of different types. On the one hand, there is direct discrimination, which is that in which categories under which it is forbidden to generate arbitrary differences, also called “suspect categories” (Castilla 2022, p. 16) – such as race, sex, gender, religion, social class, among others – are expressly used to justify unequal treatment (Morondo 2022, p. 77). In algorithmic systems, discrimination based directly on protected categories is unlikely to occur, because the way in which the information is processed implies the consideration of many elements and statistical correlations that are difficult or impossible to identify, which becomes even more difficult if the system is non-transparent (Morondo 2022, p. 77).

When data are sufficient, it is even possible to establish correlations between these data to fill gaps in information that is not available. From this information, patterns of behaviour are assumed, and decisions are made based on characteristics or behaviour, which has been called “statistical discrimination” (Williams *et al.* 2018, pp. 91–92). Given the ability to fill these information gaps, it has been argued that the prohibition of the use of protected categories is not useful, as this information can also be obtained from other data, albeit in a much more obscure and difficult to monitor form (Williams *et al.* 2018, pp. 91–92).

Regarding predictive policing systems, correlations are problematic, specifically the existence of spurious or false correlations, which connect the probability of committing a crime with variables with which they have no relation, or with characteristics that end up being discriminatory (such as race, class, age) and end up perpetuating and reinforcing that discrimination (Bennett Moses and Chan 2018, p. 811). Thus, questions arise about how these tools might be conducive to using ethnic or other profiling or promoting violations of basic criminal procedure principles, such as the presumption of innocence (Hardyns and Rummens 2018, p. 214).

In this sense, the most helpful type of discrimination for the analysis of these technologies is indirect discrimination, which focuses on outcomes – usually the only known element in the use of algorithmic systems – and determines that discrimination exists when this decision causes harm to individuals or groups of individuals belonging to protected groups subject to the “suspect categories” of discrimination (Morondo 2022, p. 77).

Biases or discrimination can occur at different points in the implementation of the system: in design and testing, during operation, or in the results (Coeckelbergh 2021, pp. 107–111). At all stages, the data used are of relevance since if they reflect biases or are unrepresentative of reality, this will be reflected in the outcome (Surden 2013, p. 4, Coeckelbergh 2021, p. 107). The quality of the result depends directly on the data quality, so the algorithm’s response will be as reliable and neutral as the data that feeds it (Mittelstadt *et al.* 2016, p. 5).

4. Women and penal system: The suspicious of feminism

The premise of legal feminism has been the realisation that law is a social creation sustained by and for men, “not neutral, but biased and saturated with values that uphold the male pattern as the standard for equality” (Cohen 2000, p. 78). For Carol Smart, the law does not necessarily act on gender as a pre-existing category but creates gender (Smart 1994, p. 177). That means that it creates what it means to be a Women, as opposed to the idea of being a Man, and within the category of women, it builds others: the bad mother, the prostitute, the criminal, etc. (Smart 1994, p. 180).

To affirm this, Smart starts from the idea that law does not follow a univocal or necessarily consistent strategy (Smart 1994, p. 180). Thus, it has been argued that it is not a simple and unitary entity but a complex and sometimes contradictory system where the representation of men’s interests does not occur uniformly (Cohen 2000, p. 95).

This position partly explains the fact that, while the law is viewed with suspicion, it is also seen as a field where a part of the feminist struggle can take advantage of the gaps left by patriarchal logic to insert its demands (Smart 1994, p. 187; Cohen 2000, p. 97). However, although the law is a form of exercising power and women can choose to make use of it, expectations of the transformative power of legal reforms should not be placed too high: it is necessary to push or promote radical proposals but mainly to be clear about the limits and how much the law can be transformed (Smart 1995, p. 129).

These ideas illustrate the complexities of using criminal law as a strategy for resolving conflicts based on structural problems of inequality, such as gender equality, and anticipate the situation that arose when the feminist movement advocated this resource to resolve problems of gender-based intimate partner violence.

One of the first reasons for suspicion against the use of criminal law is that it is insufficient to address gender issues, as it reduces a social problem to an individual issue, thereby oversimplifying and depoliticising it (Maqueda 2007, p. 398). In this sense, the response of the penal system can only be rigid, which does not allow for gradation or the possibility of revision as the conflict evolves (Pitch 2003, p. 130).

The first manifestation of this simplification is seen when constructing the criminal offence, as the need to describe the conduct precisely means that nuances and multiple interpretations must be set aside, which requires it to be precise and rigid (Pitch 2003, pp. 15–136). It is also simplified in political terms because it becomes a problem that belongs almost exclusively to the sphere of justice and in which there is no room for social or cultural issues, so in the face of the diversity of causes, it is responded to by attributing a single resolution (Pitch 2003, pp. 136–139). Finally, the criminal response generates the need to attribute responsibility to a specific individual, thus eliminating any collective responsibility derived from other spheres, such as social or cultural. (Pitch 2003, p. 137).

Nevertheless, from the 1960s onwards, part of the feminist movement opted for the penal strategy, appealing especially to its symbolic character, to make the violence suffered by women visible. However, as Bodelón (2008, p. 293) clarifies, “criminalisation was, therefore, a way of materialising the existence of the problem and making it recognisable, not the solution to it”.

One of the main consequences of the criminalisation process carried out with the interference of criminal law in matters of gender-based violence is the consequent creation of the category of victim. It is a concept that assimilates the violence suffered by women to the rest of the violence protected by criminal law, stripping it of the character of patriarchal violence and leaving women in the position of being subjects who suffer it passively, without any kind of agency (Bodelón 2008, pp. 288–289). This conception of victim incorporates a series of behavioural patterns to be followed by women, which generates an expectation for justice operators and exposes women to social sanctions when it is not fulfilled (Bodelón 2008, p. 289).

Larrauri (2008, p. 313) describes five main clichés into which women are pigeonholed: the irrational woman who withdraws the complaint, the woman who reports for financial gain, the woman who falsely reports, the woman who provokes her partner into approaching her and thereby violates the restraining order, and the vindictive woman. According to Larrauri, in most cases, these ideas stem from the imperfect and inadequate design of the criminal justice system to meet women’s needs. That results in a lack of understanding towards the complainants, as they cannot assimilate their time and reticence and only see this as a hindrance to the functioning of the judicial process. In this way, women end up being negatively labelled, for example, by blaming them for not testifying against their spouses and holding them responsible for the outcome of the trial.

This labelling affects women who face obstacles, and they begin to believe that the offence is not worth reporting. On the other hand, seeing that their needs and reasoning have no place in the system and that, even more, they are judged negatively, they begin to feel pressure to make unwanted or unsatisfactory decisions (Larrauri 2008, p. 313). In the case of Spain, it has been pointed out that one of the main problems of the legislation against gender violence (mainly the Organic Law 1/2004 of Integral Protection Measures against Gender Violence) is that it has been constructed under a security paradigm, prioritizing the punishment of the offenders, neglecting the proper protection of the victims (Heim 2017, p. 37).

The stereotypes regarding “the ideal victim” in criminal proceedings are part of a generalised problem in the legal system. In the process of interpreting and applying the law, legal operators deposit their subjectivity and beliefs about the world, acquired through socialisation in a specific culture and society. When this socialisation occurs in a patriarchal society, the embodied ideas about gender roles and gender attributes follow from this order (Custet 2021, p. 31).

This generates a bias in legal analysis, where the place of women will be understood according to the stereotype constructed in accordance with this hegemonic femininity, which is transferred to the application of law even unconsciously and generates discrimination against people who do not conform to these mandates (Custet 2021, p. 32).

5. Algorithmic tools implemented in Spain: The case of VioGén

One of the examples of predictive policing mechanisms combined with algorithms is the Integral Monitoring System for Gender Violence Cases (from now on, VioGén). It is a tool that has the function of assessing the risk of a woman reporting gender-based

intimate partner violence of being assaulted again and proposing measures for her protection. However, it also fulfils other objectives, including gathering information and facilitating its exchange between organisations with competence in gender-based violence, as well as monitoring the measures issued (González-Álvarez *et al.* 2018, p. 31).

The following section will describe the origin and functioning of the VioGén system. Using this tool as an example, an attempt will be made to show how the stereotypes of women that have been constructed in the criminal justice system (described in the previous sections) influence in concrete practice and the consequences that this may have for the protection of the right to equality and non-discrimination in the context of the use of algorithmic systems.

This analysis will focus especially on the data collection stage for the operation of the tool, for which the external audit of the VioGén system carried out by the Fundación Éticas in 2021 will be taken as a reference.

5.1. The Viogén system in perspective

This system was designed by the Secretary of State for Security of the Spanish Interior Ministry and became operational in July 2007. Its creation originates from the mandate contained in Law 1/2004, “Ley orgánica de medidas de protección integral contra la violencia de género” (LOMPVIG) (González-Álvarez *et al.* 2018, p. 30), in particular Articles 31 and 32, which are part of Title III of the law, referring to institutional guardianship, and regulate the role of the Security Forces and Corps and enshrine collaboration plans, respectively.

Specifically, the LOMPVIG created police units specialised in gender-based violence, whose function would be to carry out preventive tasks and monitor the execution of protection measures adopted in court. At the same time, the duty to generate a series of protocols to regulate the work of the police in matters of gender-based violence was enshrined, which led to the creation of the National Plan for Awareness and Prevention of Gender-Based Violence of 2007–2008, of a “Comprehensive monitoring system for cases of gender-based violence”, which is now known as the VioGén system (Sánchez López 2020, p. 123). For their part, police powers to assess the risk in which victims find themselves were incorporated into the Criminal Procedure Act in 2015,¹ due to a reform to Article 282 carried out by the Crime Victims’ Statute Act (Sánchez López 2020, p. 122).

VioGén is used throughout Spain, except in Catalonia and the Basque Country, where the police have their own risk assessment and management protocols. However, the system is interconnected with these police forces. Thus, when a victim moves to these

¹ Thus, the first paragraph of Article 282 of the Criminal Procedure Act established that the purpose of the Judicial Police is, and it shall be the duty of all its members, to investigate public crimes committed in its territory or demarcation; to carry out, according to its powers, the necessary procedures to verify them and discover the criminals, and to collect all the effects, instruments or evidence of the crime whose disappearance may be in danger, placing them at the disposal of the judicial authority. When the victims encounter the Judicial Police, they will comply with the duties of information provided for in the legislation in force. They shall also assess the particular circumstances of the victims in order to determine provisionally what protective measures should be adopted to guarantee them adequate protection without prejudice to the final decision to be taken by the Judge or Court.

places, specific protocols exist to continue dealing with the case (González-Álvarez *et al.* 2018, pp. 34–35).

Cases are initiated when they are registered in the system at the time of the complaint. A “case” is defined as a case involving a victim and a specific offender. If the same person is the victim of different offenders, there will be a case for each offender. Each case is automatically assigned to a police unit, according to the woman’s address, and this unit will be responsible for monitoring and will be the only one able to modify the information in the system (González-Álvarez *et al.* 2018, p. 33).

The case is initially assessed through the ‘Police Risk Assessment’ (VPR) questionnaire. It is followed up over time through the VPER (Police Risk Evolution Assessment) questionnaire (González-Álvarez *et al.* 2018, p. 34). The information entered in the questionnaires can have different sources, such as from the victim’s statements, from witnesses, from different police records, among others (Sánchez López 2020, p. 124).

In this way, once the VPR questionnaire has been applied, the system shows a level of risk, and for each level, specific measures are associated with it. Once the risk has been established, confirmation by the user is requested. However, although the staff member can modify this risk level – but only to increase it – in 95% of the cases the system response is confirmed (González-Álvarez *et al.* 2018, p. 37). Risk levels can be not appreciated, low, medium, high and extreme (González-Álvarez *et al.* 2020, p. 33).

In addition, a new scale (VPR5.0-H) has recently been added to the questionnaire to predict the risk of femicide. This scale has 13 indicators and estimates two levels of risk: low and high (González-Álvarez *et al.* 2020, p. 33). Therefore, as of March 2019, the VioGén form includes two scales, each with its algorithm (González-Álvarez *et al.* 2020, p. 33).

There are three types of protection measures associated with each level of risk. There are mandatory measures, others of a complementary nature and recently, victim self-protection measures have also been incorporated (Sánchez López 2020, p. 126). Each level of risk involves the incorporation of the measures of the lower risk levels (Sánchez López 2020, p. 127).

The VPER questionnaire, on the other hand, incorporates any new information or incident that could be significant to modify the consideration of the level of risk over time (González-Álvarez *et al.* 2018, p. 34). This questionnaire should be updated when requested by the judge or prosecutor, when a relevant event occurs (new complaint, breach of protection measure, among others), or ex officio by the police (Sánchez López 2020, p. 129). The periodicity of the police revision depends on the risk associated to the case. Thus, the higher the risk, the shorter the period between reviews (Sánchez López 2020, p. 130).

The first risk assessment and its modifications are communicated to the courts and the public prosecutor’s office (González-Álvarez *et al.* 2018, p. 36). Once this information has been received, it is up to the judiciary to assess the need for precautionary measures to protect the victim (Sánchez López 2020, p. 129).

5.2. *How VioGén works: Éticas Foundation Audit*

In 2021, the Éticas Foundation conducted an external audit of the VioGén² Éticas is a civil society organisation that advocates for the creation of debate and knowledge about better use of technology (Éticas Foundation, n.d.) and undertook this project because of the lack of evaluations of the VioGén system that were not conducted by the same team that created it, as well as the difficulty in obtaining information from the public administration about the functioning of this tool (Éticas Foundation 2022, pp. 5–6).

Among the issues proposed to be addressed in this audit were the transparency of the system, accountability in decision-making regarding the role of the police in validating or modifying the results of the tool, and the involvement of citizens and the people for whom the system is used in its design and monitoring (Éticas Foundation 2022, pp. 5–6).

The methodology followed was, on the one hand, to conduct a statistical analysis of Intimate Partner Homicide (IPH) cases to assess the predictive effectiveness of the system. As indicated in the report, the data used in this case are those published by the General Council of the Judiciary, which are publicly available. There is a total of 1,000 victims of intimate partner homicide (IPH), which was limited to cases occurring between 2009 and 2019, from which were excluded those in which there was no intervention of the VioGén system, that is, cases occurring in Catalonia and the Basque Country (where this system is not applied) and those in which there was no complaint (Éticas Foundation 2022, p. 20):

On the other hand, several interviews and a survey were done to study perceptions and experiences of using VioGén (Éticas Foundation 2022, p. 20). Thirty-one semi-structured telephone interviews were conducted with women whose cases were processed using VioGén system as victims of violence between 2019 and 2021 who reported the crimes in Andalusia, Valencia, Madrid or Galicia. A survey was also carried out with lawyers who are experts in the field, answered by seven professionals. Finally, interviews were conducted with two members of the Ana Bella Foundation. This civil society organisation works with women survivors of gender-based violence, which also facilitated contact with the 31 women interviewed (Éticas Foundation 2022, p. 21).

The audit results highlight various barriers to accessing the system that prevents or limit women's possibilities of reporting. Among them are the emotional barriers of the victims, derived from the situation of violence they experience; the structural barriers to which certain groups are exposed, such as those with small children, with few economic resources, belonging to rural areas, migrants, with disabilities, those who are part of the LGBT+ community, among others; and the institutional barriers, which are related to the interactions that women have with the police at the time of reporting (Éticas Foundation 2022, pp. 25–26).

² The report defines an external audit as “a process by which an independent third party examines the impact and, to the extent possible, the functioning of an algorithmic system to detect potential anomalies or practices that could be unfair or harmful towards protected groups or society. The main particularity of external algorithmic audits is that the access to the algorithm and the databases used to design, develop, test, and validate it is usually restricted” (Éticas Foundation 2022, pp. 4-5).

For the purposes of this study, the focus will be on the barriers that women encounter at the time of reporting, as at this stage they are interviewed to provide information to activate this tool and, therefore, they are the ones who have the greatest influence on data collection. The most relevant situations for these purposes that were detected in the study are: the lack of information for complainants, the emotional state they are in at the time of reporting, the type of questions included in the VioGén questionnaire, and the trust in the system due to the women's experiences in their contact with the criminal justice system.

At this stage, first, there is the problem of lack of information for the women who are going to report, who do not understand how the VioGén system works or what kind of information is expected of them (Éticas Foundation 2022, p. 26). Linked to this is the lack of legal guidance, as the majority fill in the questionnaire without first contacting a lawyer to explain the legal implications of the process (Éticas Foundation 2022, p. 27).

A second issue that comes to light is the fact that the interview is conducted at a time when most of the women are in a state of shock – as in many cases, they report the violence immediately after having suffered the episode of violence – which affects their memory and their ability to express themselves clearly and prevents them from giving precise answers (Éticas Foundation 2022, p. 27).

This situation is closely linked with the problems of the questionnaire. Both the women interviewed and lawyers express problems with the type of questions in the questionnaire, as some are ambiguous, generic or very rigid and do not allow for explanations or nuances (Éticas Foundation 2022, p. 27).

Another critical aspect of data collection is trust in the system. Indeed, as noted in the report, the level of trust in the system will affect the quality of the data that women give (Éticas Foundation 2022, p. 33). The Éticas study indicated that only 19% of women interviewed were satisfied with their experience with the system (Éticas Foundation 2022, p. 32). The negative perception of the rest of the women, whether partial or total, is based, among other reasons, on the high number of cases that are classified as “no risk” in circumstances where, for the interviewees, the mere act of reporting is risky. Another aspect that influences this perception is that physical violence is overvalued, leaving other forms of violence, such as psychological violence or cyberbullying, in the background (Éticas Foundation 2022, p. 33).

5.3. Data. Halfway between the record and the result

Based on the results of the study carried out by the Éticas Foundation, it is possible to confirm the diagnosis made by feminist currents critical of the use of criminal law to address issues of gender-based violence.

On the one hand, we observe how the system operates based on a particular type of victim and events related to the aggression. Thus, for example, the assumption that women who suffer violence assume a passive attitude towards this fact and that their participation in the solution to this conflict is no more than their collaboration with the criminal process implies that it is not considered relevant for them to have information about it, or the immediate steps to follow the complaint. That includes not having

information about how the VioGén system works and, on many occasions, not even knowing about its existence and use.

At the same time, it can be seen how the system is organised to operate based on a predominant type of violence, which is physical violence. Thus, although other types of violence can also be considered crimes, physical violence is the paradigmatic example of gender-based violence and, therefore, the one that best fits the stereotype.

On the other hand, the rigid and unidirectional nature of the response offered by the system is confirmed. That is very graphically represented in the type of questions in the questionnaire and how they will be answered without admitting degrees or nuances. The system also does not consider the emotional state of the victim, nor whether she can react in the way she is required to do so to fulfil the aims of the judicial procedure, and particularly the aims of VioGén. That is even contradictory because they are fulfilling the expectations placed on them as victims by being affected. However, this affectation prevents them from complying with the cooperation demanded by the police.

The observation of these coincidences is not trivial when it comes to circumstances that end up having a direct impact on the quality of the data with which the tool works. Thus, the design of the procedure results in the interview being carried out at a time when most of the complainants can only provide incomplete and imprecise information. In turn, the lack of understanding of how VioGén functioning prevents women from orienting their answers and to choose the most relevant information according to the objectives of this system.

These biases will directly influence the outcome, which is fundamental for the future of the women and the protection of their physical and psychological identity, as this first decision will define the protective measures taken immediately after the complaint, which often coincides with the moment of the aggression.

The most obvious, and perhaps the most serious, risk is that the system will interpret a lack of complete and sufficient information as synonymous of a low-risk level. Thus, if women are unable to provide details of the attack or to recall previous assaults accurately, the system assumes that this information does not exist and that there is, therefore, no danger.

Another possibility is that these information gaps are filled according to the perception of the police. While this type of task is always associated with a certain degree of discretion on the person, this can excessively increase when the system requires data that the complainants are not able to provide. The main danger associated with this is that those who must complete this data may do so based on the prejudices and stereotypes described by feminist criminology. There is also the risk of racist or xenophobic attitudes that associate certain cultures or nationalities with higher levels of violence, which would lead to naturalising manifestations of violence in these groups and considering them less risky.

Moving on to implementing the algorithmic system, once the data collection and recording stage is over, in which the human factor plays a leading role, it is time to evaluate the fate of these data during processing. Here, the problems and objections about algorithmic predictive policing systems described in the previous pages become particularly relevant.

Firstly, when lack of transparency does not allow external evaluations of the use of the system, it is impossible to know how the data that were so precariously collected are treated and how much weight they have in the result of the algorithm. Thus, although there is access to the questionnaire and the information requested in it, it is not known what impact each of these factors has on the final decision, leaving the question of the relevance of the answers to questions that could be very difficult for women to answer at this stage or in which there are broad levels of discretion on the part of the police.

In turn, as mentioned above, transparency is directly related to trust in the system which, at the same time affects the effectiveness of its functioning. Thus, low standards of transparency can undermine the confidence of users, which will negatively affect their outcomes and lead to an increase in mistrust.

Finally, it is essential to note the danger of discrimination that the facts described above entail. The most severe manifestation of this would be the granting of different levels of protection to complainants, depending not on the actual risk of further attacks but on their ability to tell their stories and provide information, as well as on how stereotyped they are and how prejudiced the officials who deal with them are.

6. Final reflections

This text aimed to put the spotlight on relevant issues in the use of algorithmic decision support systems in the case of gender-based violence. Not to reject the use of these technologies and promote a nostalgic return to the past, but to put into context their use, both their capabilities and their shortcomings and limitations.

In this sense, it seemed relevant to address the collection and use of data from the VioGén system because the shortcomings evident in this process could seriously affect the tool's functioning. Moreover, if we are already aware of the limitations in the capacity to anticipate future behaviours of algorithmic and predictive policing systems, an incorrect use of their logic can affect it even more.

Thus, by not being aware of these limitations, there is a risk that legal operators overestimate their capacities. That is especially relevant for those who make decisions based on this tool, but particularly for those who delegate the total weight of the decision to it.

It is also necessary to be aware of the existence of these limitations and what they consist of in order to design structures to counterbalance the possibility of making mistakes, such as, for example, the existence of subsequent review processes that do not depend solely on the data entered in the initial interview, or to encourage the use of other sources of information, hopefully immediately and prior to the adoption of the first measures. Although the protocols indicate the need for permanent reviews, it is necessary to ensure compliance with them, aware of the need to feed more and better data into the tool to optimise its functioning, and not just as part of the ritual that the algorithmic system demands of its users.

References

- Bennett Moses Moses, L., and Chan, J., 2018. Algorithmic prediction in policing: assumptions, evaluation, and accountability. *Policing and Society* [online], 28(7), 806–822. Available at: <https://doi.org/10.1080/10439463.2016.1253695>
- Bodelón, E., 2008. La violencia contra las mujeres y el derecho no-androcéntrico: pérdidas en la traducción jurídica del feminismo. In: P. Laurenzo, M. Maqueda and A. Rubio, eds., *Género, violencia y derecho*. Valencia: Tirant lo Blanch, pp. 275–299.
- Castilla, K., 2022. *Cuatro ángulos de análisis de la igualdad y la no discriminación en la inteligencia artificial* [online]. Barcelona: Institut de Drets Humans de Catalunya. Available at: <https://www.idhc.org/es/publicaciones/cuatro-angulos-de-analisis-de-la-igualdad-y-la-no-discriminacion-en-la-inteligencia-artificial.php>
- Coeckelbergh, M., 2021. *Ética de la inteligencia artificial*. 1st ed. Madrid: Cátedra.
- Cohen, B., 2000. El feminismo jurídico en los países anglosajones: el debate actual. In: H. Birgin, ed., *El derecho en el género y el género en el derecho*. 1st ed. Buenos Aires: Biblos, pp. 73–105.
- Custet, M., 2021. Argumentación jurídica y perspectiva de género: una alianza imprescindible. *Revista Derecho de Familia*, 99, 30–43.
- D'Ignazio, C., and Klein, L., 2020. *Data feminism* [online]. Cambridge, MA: The MIT Press. Available at: <https://doi.org/10.7551/mitpress/11805.001.0001>
- Éticas Foundation, 2022. *Auditoría Externa del Sistema VioGén* [online]. Barcelona. 8 March. Available at: <https://eticasfoundation.org/wp-content/uploads/2022/04/ETICAS- -Auditori%CC%81a-Externa-del-sistema-VioGe%CC%81n- -20220308.docx.pdf>
- Éticas Foundation, no date. *Why do we exist?* [online]. Available at: <https://eticasfoundation.org/about/>
- González-Álvarez, J., et al., 2018. Integral Monitoring System in Cases of Gender Violence VioGén System. *Behavior & Law Journal* [online], 4(1), 29–40. Available at: <https://doi.org/10.47442/blj.v4.i1.56>
- González-Álvarez, J.L., Santos Hermoso, J., and Camacho-Collados, M., 2020. Policía predictiva en España. Aplicación y retos futuros. *Behavior & Law Journal* [online], 6(1), 26–41. Available at: <https://doi.org/10.47442/blj.v6.i1.75>
- Hardyns, W., and Rummens, A., 2018. Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges. *European Journal on Criminal Policy and Research* [online], 24(3), 201–218. Available at: <https://doi.org/10.1007/s10610-017-9361-2>
- Harcourt, B., 2007. *Against prediction: profiling, policing, and punishing in an actuarial age*. The University of Chicago Press.
- Heim, D., 2017. Victims of gender violence: a rocky road to justice. *Rivista Di Criminologia, Vittimologia e Sicurezza* [online], 11(2), 29–42. Available at: <https://doi.org/10.14664/rcvs/723>

- Jiménez Arandia, P., 2023. *Transparencia algorítmica en el sector público* [online]. Barcelona: Generalitat de Catalunya. Available at: https://governobert.gencat.cat/web/.content/01_Que_es/recursos-i-publicacions/colleccio-govern-obert/GovernObert_9/Govern-Obert-9-cast.pdf
- Kitchin, R., 2017. Thinking critically about and researching algorithms. *Information, Communication and Society* [online], 20(1), 14–29. Available at: <https://doi.org/10.1080/1369118X.2016.1154087>
- Larrauri, E., 2008. Cinco tópicos sobre las mujeres víctimas de violencia y algunas respuestas del feminismo oficial. In: P. Laurenzo, M. Maqueda and A. Rubio, eds., *Género, violencia y derecho*. Valencia: Tirant lo Blanch, pp. 311–327.
- López Baroni, M., 2019. Las narrativas de la inteligencia artificial. *Revista de Bioética y Derecho* [online], 46, 5–28. Available at: <https://doi.org/10.1344/rbd2019.0.27280>
- Maqueda, M., 2007. ¿Es la estrategia penal una solución a la violencia contra las mujeres? *Indret* [online], 4, 475. Available at: https://indret.com/wp-content/themes/indret/pdf/475_es.pdf
- Mittelstadt, B.D., et al., 2016. The ethics of algorithms: Mapping the debate. *Big Data and Society* [online], 3(2), 1–21. Available at: <https://doi.org/10.1177/2053951716679679>
- Morondo, D., 2022. Discrimination by Machine-Based Decisions: Inputs and Limits of Anti-discrimination Law. In: B. Custers and E. Fosch-Villaronga, eds., *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice*. The Hague: T.M.C. Asser Press, pp. 73–85.
- Naffine, N., 1990. *Law and the sexes. Explorations in feminist jurisprudence*. Sydney: Allen & Unwin.
- O’Neil, C., 2017. *Armas de destrucción matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*. Madrid: Capitán Swing Libros.
- Pitch, T., 2003. *Responsabilidades limitadas. Actores, conflictos y justicia penal*. 1st ed. Buenos Aires: Ad-Hoc.
- Pitch, T., 2009. Justicia penal y libertad femenina. In: G. Nicolás and E. Bodelón, eds., *Género y dominación. Críticas feministas del derecho y el poder*. 1st ed. Barcelona: Anthropos, pp. 117–126.
- Sánchez López, B., 2020. La diligencia policial de valoración del riesgo de violencia de género en el sistema Viogén. *FORO. Revista de Ciencias Jurídicas y Sociales, Nueva Época* [online], 22(1), 119–130. Available at: <https://doi.org/10.5209/foro.66637>
- Smart, C., 1994. La mujer del discurso jurídico. In: E. Larrauri, ed., *Mujeres, derecho penal y criminología*. 1st ed. Madrid: Siglo XXI, 169–189.
- Smart, C., 1995. Legal Regulation of Male Control? In: C. Smart, *Law, Crime and Sexuality*. London: Sage, 128–145.
- Strikwerda, L., 2021. Predictive policing: The risks associated with risk assessment. *Police Journal* [online], 94(3), 422–436. Available at: <https://doi.org/10.1177/0032258X20947749>

- Surden, H., 2013. Values Embedded in Legal Artificial Intelligence. *University of Colorado Law Legal Studies Research Paper* [online], 17. Available at: <https://ssrn.com/abstract=2932333>
- Wajcman, J., 2006. *El tecnofeminismo*. Madrid: Cátedra.
- Williams, B.A., Brooks, C.F., and Shmargad, Y., 2018. How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications. *Journal of Information Policy* [online], 8, 78–115. Available at: <https://doi.org/10.5325/jinfopoli.8.2018.0078>