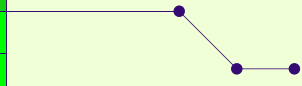
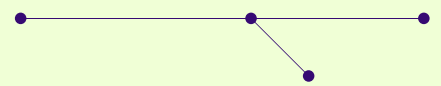


Towards a feminist  
framework for AI development:  
**from principles to practice**

Juliana Guerra



Towards a feminist  
framework for AI development:  
**from principles to practice**

Author: Juliana Guerra

English Translation: Gonzalo Bernabó

Layout and cover design: Comunas Unidas



Support By



Canada 

This work is available under a Creative Commons  
License Attribution 4.0 International (CC BY 4.0):  
<https://creativecommons.org/licenses/by/4.0/deed.es>



## Index

1. Towards a feminist framework for AI development: from principles to practice
4. Abstract
5. 1. Limitations and scope of this proposal
10. 2. Imaginaries in dispute
14. 3. All intelligence is embodied
20. 4. Facing a digestive data system
  21. 4.1. Bias
  22. 4.2. Databases
  24. 4.3. Auditability and impacts
26. 5. Doing speculatively
30. 6. Towards a guide for the development of feminist AI systems
  30. Who is this basic guide oriented to?
  30. 6.0. Before starting, understand AI as a result of collective processes.
    30. 6.1. Identify, through a collaborative process of knowledge exchange, what needs an organization or community wants to satisfy with an AI system.
    31. 6.2. Determine, based on the recognition of the context, capabilities and available resources, whether an AI system is the best way to satisfy the identified needs.
    31. 6.3. Agree on expectations, roles, timing and commitments at different stages of the development and deployment of an AI system.
    31. 6.4. Recognize, within the previously agreed working group, possible risks, opportunities and conflicts associated with the development and deployment of AI systems in their different stages.
    32. 6.5. Define, in a conscious and informed manner by the organization or community, the resources, tools and services to be used for the

## Abstract

This article is a practical approach, with a feminist perspective and located in Latin America, to the development of Artificial Intelligence (AI). Is it possible to develop AI that does not reproduce logics of oppression? To answer this question, we focus on the power relations immersed in the field of AI and make an interpretative analysis of the day-to-day experiences of seven women working in some field of AI or data science in the region, in dialogue with different statements of feminist principles and guidelines for the development and deployment of digital technologies. The narrative order of the document responds to the most relevant practical principles identified, as follows: 1. state the scope and limitations of this research; 2. how we understand the field of AI and how is the participation from Latin America in the scenarios of knowledge production in this field; 3. discrimination problems associated with the way the field of AI is currently configured and alternative proposals from feminist practices; 4. alternative proposals for the careful management of data, and to promote its auditability and reuse, from a feminist perspective; 5. practical proposals for the development of an AI that does not reproduce logics of oppression; 6. as a conclusion, a basic guide of questions and practices for AI development.

## 1. Limitations and scope of this proposal

Feminism has long ceased to be just a women's issue. Although, as Dalia de la Cerda says, the emancipation of women has not been achieved,<sup>1</sup> today we understand that colonialism, racism, classism and other historical systems of injustice continue to operate in a variety of ways and determine the life experiences - not only of women - to reproduce social structures based on the oppression of certain groups and the privilege of others.

According to the same author, oppression consists in the lack of options, and within a current feminist agenda it is necessary to be careful not to propose solutions based on privilege, but to recognize the voices, practices and knowledge that have been historically invisible, in order to continue advancing in the always unfinished project of transformation that is feminism, since "for every victory in the struggle or for every point of the agenda crossed out, there are ten that are still unfinished."<sup>2</sup>

Writing from a territory as diverse and full of contradictions as Latin America, we recognize that there is not a unique feminist project, and that much of what we collect in this text does not even carry that label. We will talk, then, about the power relations immersed in the development of artificial intelligence (AI) systems, and the possibility of developing these systems in such a way that they do not serve to continue reproducing logics of oppression.

We also recognize the limited scope of this project. In AI development, as in other areas of digital technologies and society, structural logics of oppression are manifest. For these logics to change transformations across all fields are required. There is not and never will be an AI that will by itself produce concrete transformations towards a more just society, but it is possible to change the practices of those involved in the design, production,

deployment and governance of AI in order to mitigate its detrimental impacts, and as part of a process of transformation towards a more just social life.

---

1. Dalia de la Cerda, "Feminismos sin cuarto propio," in *Tsunami 2*, ed. Gabriela Jáuregui (Mexico City: Sexto Piso, 2020), 73.

2. Dalia de la Cerda, "Feminismos sin cuarto propio", 75.

The proposals we share below are far from a solution to how different forms of oppression intersect during the development of AI systems. There is no one-size-fits-all solution, no guide with all the answers, no framework that guarantees the deployment of a fair technology. Aware of this, our effort focused on identifying and systematizing experiences of people working in the field of digital technologies in Latin America, in order to document the ways of doing and thinking about AI from our territories and from logics that are not limited to the market. We analyzed these experiences in relation to feminist principles and social justice in technologies, as we will describe below.

Moreover, our approach is not computational, although it is technical. As we will see throughout the text, for a technology such as AI, which is advancing so rapidly and currently consists of at least three branches of expertise (Natural Language Processing, Computer Vision and Robotics), it is practically impossible to imagine feminist principles embedded in a code. However, taking up the words of Kéfir, a trans feminist cooperative of free technologies, we consider that “recognizing and making visible the multiple and diverse forms of the technical re-configures our perception of who is part of a technology”<sup>3</sup> and changing such perception is key if we want to dispute the meanings of AI.

What we seek here is a description, with a feminist perspective, of the processes that take place at different moments in the development of an AI. With this, we propose practices that could influence the way the final products function and are used. The indicators for measuring the concrete impacts of these systems are beyond the scope of this text, as our proposal also implies a paradigm shift in what is considered rigorous, verifiable or sufficiently sustained in the field of digital technologies, which includes AI.

Paz Peña and Joana Varon have developed a framework for analyzing potential harms of algorithmic decision making by holistically considering the power relations embedded in the deployment of AI systems by the public sector. For them, “instead of asking how to develop and deploy an A.I. system, shouldn’t we be asking first “why to build it?”, “is it really needed?”, “on whose request?”, “who profits?”, “who loses?” from the deployment of a particular A.I. system? Should it even be developed and deployed?”<sup>4</sup>

---

3. Kéfir, “Guía de migración de infraestructura”, 2020, <https://archive.org/details/guia-migracion-Infraestructura>, 10.

4. Paz Peña and Joana Varón, “Oppressive A.I.: Feminist Categories to Understand its Political Effects”, Not my A.I., October 10, 2021, <https://notmy.ai/news/oppressive-a-i-feminist-categories-to-understand-its-political-effects/>

This approach is similar to that adopted by Patricio Velasco and Jamila Venturini in their analysis of AI implementation by the public sector in Latin America. Based on four case studies, they conclude that there is little provision in the public sector to evaluate the relevance and necessity of the proposed technological solutions, and even less to identify the demands and proposals of the people potentially affected in their design. In addition, they draw attention to the fact that “any deployment of technology takes place in a space with political tensions, where the danger is in pretending that the systems considered may ignore, hide or offset these tensions with nothing more than the very claims of efficiency underpinning them.”<sup>5</sup>

In the face of these important findings and questions, we seek to bring feminist practices closer to those involved in the development, deployment and governance of AI systems. We do so in response to the call made years ago by Safiya Umoja Noble, to “keep sufficient feminist pressure on the development of technologies, in the context of material consequences that diminish any liberatory possibility.”<sup>6</sup>

To elaborate this proposal we made a non-exhaustive review of different audit frameworks and impact assessments of algorithmic systems, in relation to human rights, seeking to understand what kind of mechanisms are being proposed in practice. The references on these issues are mainly in English and located in the United States, which reproduces global power relations already identified by other studies.<sup>7</sup>

We review six statements of feminist principles or values, which have emerged from collective reflection during workshops or longer-term processes, and are situated in diverse contexts: The Feminist Principles

---

5. Patricio Velasco and Jamila Venturini, “Automated decision-making in public administration in Latin America”, March 2021, [https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC\\_informeComparado.pdf](https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informeComparado.pdf), 32.

6. Safiya Umoja Noble, “Traversing Technologies. A Future for Intersectional Black Feminist Technology Studies”, *The Scholar & Feminist Online Issue* 13.3 -14.1 (2016), <https://sfonline.barnard.edu/traversing-technologies/safiya-umoja-noble-a-future-for-intersectional-black-feminist-technology-studies>.

7. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, vol. 1 n. 9, 389–399. Retrieved on February 22, 2021 from <https://www.nature.com/articles/s42256-019-0088-2.pdf>.

of the Internet,<sup>8</sup> The Oracle for Transfeminist Technologies,<sup>9</sup> Manifest-No,<sup>10</sup> Design Justice Principles,<sup>11</sup> the Feminist Principles for the TICas Program<sup>12</sup> and AI Decolonial Manyfesto.<sup>13</sup> We also reviewed the Data Feminism Principles,<sup>14</sup> and the Feminist Categories to Understand [AI] Political Effects,<sup>15</sup> resulting from the research carried out by their authors.

We examined the practical proposals of three recent feminist guides for the deployment and management of different digital technologies, produced in Latin America: *Guía de Migración de Infraestructura* (Infrastructure Migration Guide) by Kéfir cooperative,<sup>16</sup> *Guia para aprendizaje e construção de redes comunitárias* (Guide for learning and building community networks) by Marialab<sup>17</sup> association and the methodological guide for the research on Feminism, Ethics and Geospatial Data, conducted by Selene Yang together with ILDA Foundation.<sup>18</sup>

Finally, we interviewed seven women working in some branch of AI or data science in Latin America. Through their words and experiences we organize the information consulted, and propose a preliminary

---

8. Feminist Principles of the Internet <https://feministinternet.org/>.

9. The Oracle for Transfeminist Technologies <https://www.transfeministech.codingrights.org>

10. Feminist Data Manifest-No <https://www.manifestno.com>

11. Design Justice Network Principles <https://designjustice.org>

12. TICas Program. Feminist Principles, CitizenLab Summer Institute, 2019.

13. AI Decolonial Manyfesto <https://manyfesto.ai/index1.html>

14. Catherine D'Ignazio and Lauren F. Klein, *Data Feminism* (Cambridge: MIT Press, 2020). <https://data-feminism.mitpress.mit.edu>

15. Peña and Varón, "Oppressive A.I."

16. Kéfir, "Guía de migración de infraestructura".

17. Carla Jancz, "Enredando territórios de cuidado: Guia para aprendizaje e construção de redes comunitárias", Marialab, 2021. <https://www.marialab.org/wp-content/uploads/2021/03/Cartilha-de-redes-comunitarias-FINAL.pdf>.

18. Selene Yang. "Feminismo, ética y datos geoespaciales: guía práctica de recomendaciones", ILDA, work documents. <https://doi.org/10.5281/zenodo.4913074>.



practical guide that gathers feminist concerns and stakes in different moments of AI development, deployment and governance. They are:

- Ximena Gutiérrez, postdoctoral researcher in the area of Natural Language Processing (NLP) at University of Zurich, and part of Elotl community.<sup>19</sup>
- Marisol Flores Garrido, professor and researcher at the Bachelor's Degree in Information Technologies in Science, Universidad Nacional Autónoma de México - Morelia.
- Sofía Trejo, PhD in Mathematics, teacher and independent researcher in Artificial Intelligence ethics.
- Luciana Benotti, Professor at the Department of Computer Science and researcher in the area of Natural Language Processing (NLP), Universidad Nacional de Córdoba, Argentina.
- Selene Yang, PhD candidate at the Communications Program, Universidad Nacional de La Plata, Argentina. Co-founder and coordinator of Geo Chicas<sup>20</sup> within the Open Street Map community.<sup>21</sup>
- Fernanda Carles, mechatronics engineer and independent data science researcher in Paraguay.
- Carolina Padilla, master's degree in Analytical Intelligence and current leader in advanced analytics at an advertising company, Colombia.

For a more fluent reading, and considering we only conducted one interview with each person, from now on, we will name the interviewees each time we quote their words, but we will only make the citation in a footnote the first time.

---

19. Comunidad Elotl. Technologies for Mexican Languages <https://elotl.mx>

20. Selene Yang, Céline Jacquin y Miriam González, "Geochicas: Helping Women Find their Place on the Map", Mapillary, Mayo 28, 2019, <https://blog.mapillary.com/update/2019/05/28/putting-women-on-the-map-with-geochicas.html>

21. OpenStreetMap <https://www.openstreetmap.org>

## 2. Imaginaries in dispute

AI can be many things. For the purposes of this text, we will not go into the detail of what defines the field, its branches and technical characteristics, as these have been investigated and exposed elsewhere before, from which we take five relevant points for a feminist perspective. First, today the term Artificial Intelligence is used much more by the industry for promotion and marketing, while in computational technical research it is referred to as “machine learning”.<sup>22</sup>

Second, AI-based computational systems are intended to make decisions, and those decisions are expected to have real-world effects. Third, AI developments in recent years are based on neural network systems, also called “deep learning” systems. These recent advances are made possible by three conditions: a consolidation of the field of machine learning; the massive availability of data; and the increase in computational capacity on that data (i.e., hardware and processing).<sup>23</sup>

Fourth, in machine learning systems, the first source of value comes from the data to be processed, and the second from the algorithms that process these data. In the processing of the data, statistical models are defined, which are the ones that will generate a result (for example, a decision). However, the data for learning are not raw, but have gone through a previous work of cleaning, organizing, formatting and labeling. And it is in these previous processes that the final results are determined, much more than in the algorithmic parameters.<sup>24</sup>

Fifth, AI is much more than a matter of computation, machines and data. In Kate Crawford's words, AI is “inherently political” because it is being permanently shaped by a set of technical and social practices, as well as infrastructures, institutions and norms, and it is also material because it is composed of natural resources, energy and human labor.<sup>25</sup>

---

22. Kate Crawford, *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale University Press, 2021).

23. Iván Meza Ruiz, Sofía Trejo and Fernanda López, “¿Quién controla los sistemas de Inteligencia Artificial?” (Who controls Artificial Intelligence systems?), *Revista Anfibia*, June 1, 2021, <https://www.revistaanfibia.com/quien-controla-los-sistemas-inteligencia-artificial/>

24. Vladan Joler y Matteo Pasquinelli, “The Noosphere Manifested. AI as Instrument of Knowledge Extractivism” *Noosphere*, 2020, <https://noosphere.ai>

25. Crawford, *Atlas of AI*, 8.

As Decolonial Manyfesto states, “AI is a technology, a science, a business, a knowledge system, a set of narratives, of relationships, an imaginary.”<sup>26</sup>

With these five points in hand, we will from now on speak of AI to refer to neural network or “deep learning” systems (not other areas of “machine learning”) applied to technologies that may be deployed by public or private actors, and with which many of us interact daily. Furthermore, considering power relations in AI, we will also refer to the materiality of these systems, including the datasets that feed them.

Much of the power mobilized through AI is given because, while we increasingly interact with biometric systems, voice assistants or word auto-completion, we know very little about how these systems work, who builds, powers and analyzes them, what is required for them to exist, and who benefits from their use. For a person developing systems in Latin America, moreover, it is not enough to understand the operating logic of any of these systems, to deploy them at will and in the context that best suits their interests. It requires, at least, access to large amounts of data, ideally organized and labeled for the purposes being pursued, and robust computing equipment.

According to Luciana Benotti, most AI technologies are developed in the United States, usually by men. And so that they do not only benefit themselves, “it is necessary for a diverse group of people to be involved in the development of these tools at different levels.” She does research in the area of Natural Language Processing because, she says, although tools that use this technology are widely used in Latin America, such as Google searches, automatic translation or text auto-completion, “there is very little participation and very little representation of our languages and accents.”

For her, the main difference, “at the academic level in universities, is that in the northern hemisphere there are many more, there is much more money in this area. [...] In Argentina, the area of computing that includes Artificial Intelligence is very small, there are very few researchers and it is being reduced due to an economic issue, because people are going abroad or to the industry.”<sup>27</sup>

---

26. AI Decolonial Manyfesto.

27. Luciana Benotti (NLP- Academy) author interview, April 2022.

A similar perception is shared by Mexican researcher Ximena Gutiérrez, from Zurich. “The Latin American community and the Mexican community is underrepresented in the major conferences in the area,” she says. For her, “you have much greater visibility when you are at a university in the Global North, Europe or the United States.” And she recognizes as problematic that the system of academic incentives in Mexico fails to value participation in conferences, “but ironically, because of the speed of machine learning, everything is published at conferences. You cannot wait to publish something in two or three years because it is considered old, you have to be at the conferences because that is where everything moves.”<sup>28</sup>

Regarding conferences, Luciana Benotti comments that now “they are held more often and have ten times the number of people they had a few years ago or more. What has also changed is where people come from. [...] The number of papers from companies has increased, the number of papers coming from China has increased and in the northern hemisphere it is as if university and company research is almost indistinguishable. In the companies there are people almost only doing research and then this is integrated to the products, it is like the cutting edge research of Artificial Intelligence is also in the industry.”

About research funding in Mexico, Marisol Flores questions “the models that are imposed in the famous literature, which do not dialogue with the reality of the country. It bothers me a lot the issue of autonomous cars, where I see that there is a very large investment. I feel that in my country they do not solve any problem, not everyone is going to have them, and all that is in the head of the creators is this European-gringo male fantasy of a car that drives itself.”<sup>29</sup>

The participation of people from the global south in AI (and computer science in general) knowledge development and construction spaces is still very low, as well as the participation of women, LGBTTTQI+ people, racialized people and people with functional diversity. For this reason, more and more conferences such as ICML,<sup>30</sup> among others, have “Diversity and Inclusion” programs, which include events, tracks and specific projects for underrepresented communities, as well as funding programs for physical and remote participation, and the implementation of Codes of Conduct or Anti-Harassment Policies.

---

28. Ximena Gutiérrez (NLP-Academy) author interview, April 2022.

29. Marisol Flores Garrido (Research, communication-Academia) author interview, April 2022.

30. International Conference on Machine Learning | <https://icml.cc/public/DiversityInclusion>

The Association for Computational Linguistics has participation statistics from its 2017 annual conference,<sup>31</sup> which discriminate female participation (24%), male (70%), and “other/not stated/unknown” (6%), as well as participation by region, “North, Central and South America” (55%), “Europe, Middle East and Africa” (19%), “Asia, Pacific” (23%), and “Unknown” (3%). These figures speak little of regional differences, for example North-South, and also of other diversities.

Projects such as the Oracle for Transfeminist Technologies seek to account for this situation and to be inspired to imagine what kind of values could be embedded in AI systems if they were built from other subjects and in other spaces. As stated on their website “the wisdom of the oracle, embedded with transfeminist values, will help us foresee a future where technologies are designed by people who are too often excluded from or targeted by technology in today's world.”<sup>32</sup>

We will return later to how the imagination has been a powerful tool of dispute in the face of imposed models of technological development. As Donna Haraway stated some time ago in her *Cyborg Manifesto*, one of the most influential texts within the feminist currents of thought and critique of technology, “liberation rests on the construction of the consciousness, the imaginative apprehension, of oppression, and so of possibility.”<sup>33</sup>

---

31. ACL Diversity Statistics <https://www.aclweb.org/portal/content/acl-diversity-statistics>

32. The Oracle for Transfeminist Technologies <https://www.transfeministech.codingrights.org>

33. Donna Haraway. “A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century”, in *Simians, Cyborgs and Women: The Reinvention of Nature*, (New York: Routledge, 1991), 149.

### 3. All intelligence is embodied

What people are excluded from technology, and who are targets of technology? Contradictory as it may seem, it is usually the same groups of people. Many of the discriminatory effects of AI are attributed to the interests that motivate its design and deployment, the people and institutions involved in this process, and the contexts where it is developed, as opposed to the contexts where it is deployed and the people to whom it is directed.<sup>34</sup> As Jamila Venturini and María Paz Canales point out,

“with the idea of digital transformation gaining steam in Latin America, several data entry points have been created to collect information that map traditionally marginalized groups who are insufficiently understood or prioritized in welfare policies. The problem is that while doing so, more often than not, states mistake data for bodies. With the same racist and patriarchal foundations that were responsible for making Latin America one of the most unequal regions of the world, such initiatives risk creating new and worse forms of control if they are silent on the biases that are encoded in the way data is generated, collected, and used to build technical systems and narratives that contribute to decision making.”<sup>35</sup>

But this is not just in the public sector. As we saw earlier, initiatives to include underrepresented groups in the development of AI technologies reflect the lack of diversity in these spaces, which redounds to the paradigms on which the field has developed. As Catherine D'Ignazio and Lauren Klein point out on the visualization of statistical information, one of the main reasoning in data science is that “the more plain, the more neutral; the more neutral, the more objective; and the more objective, the more true.” However, they say, this is only a veracity effect or, as Donna Haraway called it some time ago, “the *god trick*: it’s a trick because it makes the viewer believe that they can see everything, all at once, from an imaginary and impossible standpoint. But it’s also a trick because what appears to be everything, and what appears to be neutral, is always what she terms a *partial perspective*.”<sup>36</sup>

---

34. Ruha Benjamin, *Race after Technology. Abolitionist tools for the New Jim Code*, (Cambridge:Polity Press, 2019).

35. Jamila Venturini y María Paz Canales, “A Feminist Lead Towards an Alternative Digital Future for Latin America” Bot Populi, April 20, 2022, <https://botpopuli.net/a-feminist-lead-towards-an-alternative-digital-future-for-latin-america>

36. D'Ignazio y Klein, *Data Feminism*, 76. (traducción propia)

In this context, it is the one-sided perspective of a very homogeneous group of people, mainly white males in the United States, which is overshadowed by the ideas of neutrality, objectivity and truthfulness on which AI systems and digital technologies in general are developed. Or, as different authors call it: models, narratives and culture,<sup>37</sup> or the vision,<sup>38</sup> or the ideology<sup>39</sup> of Silicon Valley being deployed in various other contexts to respond to structural forms of oppression that, most of the time, they do not understand because they are far removed from their daily realities.

And as Ruha Benjamin has pointed out, the problem is not that developers in Silicon Valley deliberately choose to reproduce oppression on underrepresented groups, but rather that “power operates at the level of institutions and individuals - our political and mental structures- shaping citizen-subjects who prioritize efficiency over equity.”<sup>40</sup> In the context of Latin America, for example, it is not only a matter of hiring the services of companies located in Silicon Valley, but of reproducing their models, reusing their codes and databases, considering efficiency under the argument that we have few resources to produce localized systems based on a critical reading of our own contexts.

Fernanda Carles says that she once had “the opportunity to work as a technology mentor for a startup that won a competition to make an application to tackle violence against women.” When she reviewed the project, she noticed that it was very invasive because it collected and stored too much information, and for this and other reasons they had to “completely change the application.”<sup>41</sup> Marisol Flores says that years ago a group of women in Mexico won a feminist artificial intelligence award. The project consisted of “a bot that was supposed to detect whether you were in danger of violence or not.” She says that while she welcomes the initiative, she has

“problems with how the public's imagination is being sold. A bot that is going to predict if your body is going to be exposed seems super

---

37. Venturini and Canales, “A Feminist Lead”.

38. Joana Varon and Paz Peña, “Artificial intelligence and consent: a feminist anti-colonial critique”, *Internet Policy Review*, 10(4) (2021). <https://doi.org/10.14763/2021.4.1602>.

39. Yásnaya Elena Aguilar Gil, “A modest proposal to save the world”, *Rest of World*, December 9, 2020, <https://restofworld.org/2020/tecnologia-tequio-cambio-climatico>

40. Benjamin, *Race after Technology*, 30-31.

41. Fernanda Carles (Neural Networks-Academia, Civil Society) author interview, April 2022.

absurd to me. [...] What I think is that since they are trying to make women more empowered, this also lends itself to confusion and there are a lot of useless tools just to say that women are involved.”

These are not the only app ideas to protect women from gender-based violence, nor are they the only ones that fail to understand the needs a woman may have when she is facing violence.<sup>42</sup> And there are many other examples of technologies that, by design, are far from contributing to more equitable relationships. Selene Yang says that “there are many countries in the world where, in public restrooms, there are condom vending machines in men's restrooms and pad and tampon vending machines in women's restrooms.” That, she says, speaks to the fact that access to reproductive rights is given merely to men. But

“when you look it up on an OpenStreetMap map it tells you there are 10 feminine pads vending machines, and 1970 condom machines. That means there are more men mapping. That tells us a lot of things: the masculinization of spaces, the lack of representation of women mapping their interests, the issue of the binarity of space. What does this mean for men who menstruate or women who fuck heterosexually?”<sup>43</sup>

If we consider that data is the primary source of value for machine learning systems, who captures, labels and defines the rules for processing it is a central issue for thinking about AI systems that do not reproduce oppressive logics. Carolina Padilla says that in working with data “the number of women who participate, whether in leading positions or being in projects, is always very small.” In her experience, whenever they make a job offer, “80% of those who apply are men. Women do not apply for data or technology areas.” In addition, “usually women are in the commercial teams, in marketing, but in hard topics such as data, which involve convincing or explaining things that not everyone understands, then you practically have to go to high ranks where men usually lead.”<sup>44</sup>

These examples are illustrative of two increasingly visible structural conditions: the digital gender gap and violence against women. But there are other forms of oppression in the field of digital technologies that we tend to see less. For example, women and men working in

---

42. Marianne Díaz “Eliminar la violencia de género con un clic”, Derechos Digitales, January 12, 2017, <https://www.derechosdigitales.org/10730/eliminar-la-violencia-de-genero-clic>

43. Selene Yang (Data Science-Community) author interview, April 2022.

44. Carolina Padilla (Data Analytics - Industry) author interview, April 2022.



assembly factories for the Mexican electronics industry, with clients as large as Tesla, Nokia, IBM or Foxconn, demand dignified working conditions such as spaces free of harassment and discrimination, and minimum health conditions for handling highly toxic materials. In addition, they demand that companies be consistent with the social responsibility seals they display on their websites, as they maintain unfair and often illegal working conditions.<sup>45</sup> These and other activities part of the AI production chain,<sup>46</sup> configure and sustain a system of oppression that can hardly be corrected with computer codes alone.

In another area, Ximena Gutiérrez participates in a community that seeks to reduce bias through the production of tools and digital content in indigenous languages from Mexico. She recognizes that

“making technology for all these languages is not just a matter of saying “ok I'm going to put everything online” but realizing that these bias come from structural inequalities. In Mexico, and I would dare say in Latin America as well, if you are born speaking an indigenous language your chances are going to be very different from someone who does not speak an indigenous language and is white.”

In the words of Mixe linguist Yásnaya Aguilar, “being bilingual is not the same as being bilingual.”<sup>47</sup> Specifically, in Mexico, speaking an indigenous language and Spanish does not have the same social status as speaking Spanish and English. And just as bilingualism is relative, so is the place from which it is viewed. Regarding the technology industry, the author mentions that

“to metropolitan eyes, Latin America, with its minimal patent production and negligible investment in science and technology, lags behind. The Silicon Valleys springing up in different parts of the world - more an ideological concept than a geographic location- have long dismissed the region as a passive receptor for technology. But Abya Yala [a term the Guna people of Panama use to describe the Western Hemisphere's largest landmass as it existed before 1492] challenges that narrative. Here on the periphery, technology, when repurposed for resistance, can bolster the autonomy of peoples and communities”.<sup>48</sup>

---

45. Cetien, *Voces obreras / Workers' Voices*, (Guadalajara: Rosa Luxemburgo Stiftung, 2018), [https://www.rosalux.org.mx/sites/default/files/voces\\_obreras-cetien-version2019.pdf](https://www.rosalux.org.mx/sites/default/files/voces_obreras-cetien-version2019.pdf)

46. Crawford, *Atlas of AI*.

47. A. Gil, *Ää: manifiestos sobre la diversidad lingüística*, (Mexico City: Almadía, 2020), 31.

48. Aguilar Gil, “A modest proposal to save the world”.

This challenge is precisely what gives meaning to this text. For years, feminist groups have been organizing in Latin America to build alternatives<sup>49</sup> to the violent and oppressive order that prevails in digital technologies. In the words of the Manifesto for hackfeminist algorithms, “we want to resist against all infrastructure that enables and reproduces oppression, discrimination and misogyny, through our bodies-territories-algorithms in whatever space we inhabit on or off the Internet.”<sup>50</sup>

For this resistance, infrastructure is a central issue. That is, in addition to the question of what data, and whose data, is being used in digital technologies, the place where that data is hosted matters, as does where that place is, who owns it, who has access to it and under what conditions. But also, how the system of data capture and traffic is kept permanently running. As stated in the Guide for learning and building community networks, people who use the Internet often do not have a concrete knowledge of its infrastructure. “This information is hidden behind the idea of a “cloud” where everything seems very abstract and magical.” But the infrastructure is more like a network of terrestrial and submarine cables, behind which “there are very robust routers controlled by companies and governments.”<sup>51</sup>

Conceiving digital infrastructures as “territories of care” is another way of recognizing that technologies also have to do with the ways we relate to each other, beyond the digital tools we are developing, deploying or using.<sup>52</sup> In Selene Yang's words, it is important to recognize, for example, “academic housework. This means that what women and dissident genders do in the panels and conferences is to organize them, organize the press release, the logo, the food and all that. It is a lot of work of organization, care and coordination.”

Making visible, re-signifying and putting care work at the center has been part of feminist bets, also in the field of digital technologies. It is what allows us to see that AI is not only made of data and algorithms,

---

49. Derechos Digitales, ed., “Latin America in a Glimpse. Gender, Feminism and the Internet” (Derechos Digitales, APC: 2017), [https://www.derechosdigitales.org/wp-content/uploads/GlImpse2017\\_eng.pdf](https://www.derechosdigitales.org/wp-content/uploads/GlImpse2017_eng.pdf)

50. Liliana Zaragoza Cano and Anna Akhmatova, “Manifiesto por algoritmias hackfeministas”, GenderIT, october 15, 2018, <https://www.genderit.org/es/articles/edicion-especial-manifiesto-por-algoritmias-hackfeministas>.

51. Jancz, “Enredando territórios de cuidado”, 24-25.

52. Jancz, “Enredando territórios de cuidado”.

that there is a whole materiality of work that is often feminized, precarious and hidden. In order not to reproduce logics of oppression, it is necessary to contemplate these jobs as well. As Paola Ricaurte states,

“it is essential that we listen to and learn from the embodied experiences of datafication, algorithmic mediation, and automation in the lives of women and girls, indigenous communities, immigrants, refugees, platform workers, non-binary people, and rural communities across the globe. Their situated knowledge(s) can help us understand algorithmic harms as a more complex phenomenon than a straightforward experience.”<sup>53</sup>

---

53. Paola Ricaurte, “Artificial Intelligence and the Feminist Decolonial Imagination”, March 2, 2022. <https://botpopuli.net/artificial-intelligence-and-the-feminist-decolonial-imagination/>.

## 4. Facing a digestive data system

To summarize in one sentence what, from her perspective, AI would require to be feminist, Fernanda Carles reaffirms: “it has to serve humanity and not the market, not productivity.” This resonates with the words of Sofía Trejo, for whom “a lot of the criticism goes towards big corporations and the way they are developing these [AI] systems, but they are very much related to new ways of creating resources and there is a whole new form of data commerce.”<sup>54</sup>

For a person connected to the Internet today, it is difficult to identify how many AI systems they interacts with on a daily basis, for example when using the basic functionalities of a cell phone, such as screen sensors or virtual keyboards. On the other hand, more and more states are implementing algorithmic systems for the provision of social services,<sup>55</sup> and in these cases, it is also often difficult to identify which entities are in charge of data collection and processing. According to Joana Varon and Paz Peña,

“in Latin America, IBM, Microsoft, NEC, Cisco, Google are commonly involved in AI projects developed by the public sector from the region. Every project feeds databases and provides intelligence for machine learning systems of these companies, which can use these less regulated environments, where enforcement of privacy rights is weak, as laboratories to test and improve their systems, normally unaccountable to possible harmful consequences.”

Therefore, the possibility of conceiving an AI outside the market seems distant. However, as Yásnaya Aguilar says and as we mentioned before, narratives are also a terrain of struggle, and in terms of data and digital technologies, the narrative “the all-encompassing Western mythology flooding our distant territories” is not only being questioned but also challenged with different initiatives of resistance and collaborative work. In this part, we collect alternative proposals for the careful management of data, and to promote its auditability and reusability, from a feminist perspective.<sup>56</sup>

Therefore, the possibility of conceiving an AI outside the market seems distant. However, as Yásnaya Aguilar says and as we mentioned

---

54. Sofia Trejo, (AI Ethics - Academia, freelance) interview with the author, April 2022.

55. This has been documented at <https://ia.derechosdigitales.org> y <https://notmy.ai/es>

56. Varon and Peña, “Artificial intelligence and consent.

before, narratives are also a terrain of struggle, and in terms of data and digital technologies, the narrative “the all-encompassing Western mythology flooding our distant territories”<sup>57</sup> is not only being questioned but also challenged with different initiatives of resistance and collaborative work. In this part, we collect alternative proposals for the careful management of data, and to promote its auditability and reusability, from a feminist perspective.

#### **4.1. Bias**

One of the main problems recognized today in AI systems is bias. In the United States, where the largest corporations are located, a growing field of algorithmic audits has emerged, to identify and propose ways to mitigate bias and harms. More recently, the implementation of algorithmic impact assessments is being promoted, taking as a reference impact assessment models in other areas - going beyond the technical audit on algorithms- and regulation is also being pushed to determine accountability standards for AI companies.<sup>58</sup>

However, as some authors have criticized, the emphasis on biases and algorithmic accountability falls short of understanding how oppression is reinforced and reproduced in AI systems,<sup>59</sup> and in many cases is helping companies profit from “bias minimization” labels in their products.<sup>60</sup> Therefore, from Data Feminism they propose to understand and design systems that deal with the origins of biases, that is, structural oppression.<sup>61</sup> Referring to the practice of mapping, Selene Yang argues that

“biases can be mitigated, but more than that, they must be problematized. For me, the horizon and the problematization itself is the recognition of biases, it is to understand where it comes from in order to then operate around the bias. You will not be able to remove people's

---

57. Aguilar Gil, “A modest proposal”.

58. Emanuel Moss et al., “Assembling Accountability. Algorithmic Impact Assessment for the Public Interest” (Data & Society, 2021), <https://datasociety.net/wp-content/uploads/2021/06/Assembling-Accountability.pdf>.

59. Paz Peña y Joana Varón, “Oppressive A.I.”

60. Benjamin, *Race after Technology*.

61. Catherine D'Ignazio y Lauren F. Klein, *Data Feminism*.

bias, you can educate, discuss and dialogue with people, but many times the bias will not go away. There can be very implicit bias or explicit bias. I think explicit bias is more workable because it is stated.”

Sofía Trejo has a similar perspective, as she considers that “it is very useful not to try to remove biases but to use them to understand the discrimination problems we have, to understand the way these types of discrimination are being reinforced.” She believes that focusing too much on biases “is to point to solutions that do not fix the system itself, but rather create a new justification that they exist to start something that is fairer. But that justice to which we want to aim is not built on democracy and consensus.” And as Marisol Flores points out, “sometimes the biases are conflicts of value, such as, for example, whether I prefer privacy or equity in certain things.”

In order to make biases explicit, it is necessary not only to develop technical tools for their identification, but also to open dialogues on the subject and agree on data access policies. Luciana Benotti says that, in her experience, today the relationship between industry and academia is very asymmetrical because companies offer students to do research in exchange for access to data, that is, without financial remuneration, which implies a double exploitation: on the one hand, data extraction and, on the other, the precariousness of work. But also because it is very difficult for the products of such research to be used later. “If the company does not release the data it is irreproducible work and to me that is not science. For me, science is making enough information available so that others can reproduce it. Anything else is propaganda.”

#### **4.2. Databases**

For machine learning research, nowadays, large benchmark datasets are available on the Internet to be used and adapted by different research communities. As pointed out in a recent article, the different research tasks are concentrated in an increasingly reduced number of datasets, many of which are in the hands of fewer and fewer institutions (companies such as Microsoft and Google, universities such as Stanford and Princeton). In addition, a database can be used for tasks very different from those for which it was designed.<sup>62</sup> It is worth noting that, due to the access conditions imposed by their owners, although these datasets can be used and adapted, any bias mitigation work done on them does not guarantee an improvement in the datasets. On the contrary, this situation implies that different implementations around

---

62. Bernard Koch et al., “Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research,” *ArXiv: 2112.01716 [Cs, Stat]*, December 3, 2021, <http://arxiv.org/abs/2112.01716>

the world are relying on models trained by datasets whose collection and labeling policies, to say the least, are not entirely transparent.

For Marisol Flores, from a feminist perspective it is very important to “question the trust we have in data” and that is why it is necessary to document the data collection processes. She recognizes, however, that this careful work with data is possible because she has a full-time job. “It would be easier for my career to download a dataset from another country and test the technical tools, than to be looking for who has information on footbridges. I think what we pay with is time, and time is very valuable.”

For the work with geospatial data, from a feminist perspective it has been recognized the importance of having standardized formats for data capture and management, and thus promote the interoperability of databases and applications. Also, documenting processes to enable the reproducibility of both data and findings.<sup>63</sup> But documentation also serves to attribute and make visible the work done. As we have shown so far, in AI there are many jobs that are not considered as such when it comes to emphasizing algorithms and data. And although it is these elements that differentiate machine learning from other fields of digital technologies, care work is an integral part of any productive activity. Finally, in view of the participation gaps we have briefly outlined here, it is very important to recognize who and how performs the different tasks that make possible the development of AI systems or any other technology.

Regarding policies for data use, and as a reference, we consider it relevant to draw attention to the Feminist Peer Production License (f2f), as a political statement of the way data is expected to be reused. This license was proposed by the feminist hacklab *la\_bekka* in October 2019 for their fanzine on feminist servers, and states:

This fanzine is released under the Feminist Peer Production license. This means that, recognizing the authorship of the work, you are allowed to share it (copy it, distribute it, perform it and communicate it publicly) and make derivative works under the following conditions: exploit it commercially only if you are part of a cooperative, non-profit organization or collective, or organization of self-managed female workers, who defend and organize under feminist principles; and, share it with the same license.<sup>64</sup>

---

63. Selene Yang, “Feminismo, ética y datos geoespaciales”.

64. Inés Binder and m4rtu, “Cómo montar una servidora feminista con una conexión casera”, (“The right to have our own infrastructure: how to set up a feminist server with a home connection”) *la\_bekka*, october 9, 2019, <https://labekka.red/servidoras-feministas/>

It is important to differentiate between two types of datasets, for their treatment and use from a feminist perspective. On the one hand, datasets created in a process based on feminist principles, and on the other hand, datasets already existent in whose design there was no participation. The latter, as has been mentioned in so many places, are a record of the past and using them without this awareness to train and feed automatic learning systems is continuing to reproduce labels, stereotypes and roles, from particular (and privileged) points of view just as the racist,<sup>65</sup> misogynist, transphobic and classist order we still live in has been configured.

As several interviewees pointed out in this research, identifying biases is also the possibility of identifying problems such as different forms of discrimination or other issues that feminist practice seeks to transform. And the machine learning tools we know today can serve these purposes, for example to identify patterns of spatial behavior of the forced disappearance of women in a period of time,<sup>66</sup> considering that this is a current problem, concrete, with little visibility and with enormous effects on social life.

We will refer to the datasets that are built within processes with feminist principles in the next part. Before that, we will make some points regarding the auditability and impact assessment of AI systems, understanding that power is also structured in these processes, often reinforcing oppressive orders, but also opening space to contest them.<sup>67</sup>

### **4.3. Auditability and impacts**

As Data & Society's analysis based on the U.S. context points out, while the existence of algorithmic impact assessment frameworks provides a public space for scrutiny, which gives legitimacy to the accountability process, some of the existing frameworks, such as human rights and privacy impact assessments, tend to be opaque, either because of the legal regime that covers them, in the first case, or because of the technical complexity of the documents, in the second case.<sup>68</sup> From a feminist perspective, the replicability objective expected with process documentation is also to make them readable, that is, to use interfaces,

---

65. Benjamin, *Race after Technology*.

66. This example was shared by Marisol Flores Garrido, in an interview with the author in April 2022. Due to the sensitivity of the subject we do not share the textual quote but we consider it very relevant, as we explain in the body of the text.

67. Emanuel Moss et al., "Assembling Accountability".

68. Emanuel Moss et al., "Assembling Accountability".



formats and languages that are accessible to the people who are expected to use such documentation.

This goes beyond the right to explainability of AI systems, as recognized by the European General Data Protection Regulation,<sup>69</sup> and also goes beyond proposals to include communities in the algorithmic accountability process by recognizing their experiences as a form of expertise, and thus, as authoritative voices to participate in evaluations.<sup>70</sup> As we will see in the next part, the people to whom the documentation is addressed are not only considered users of the systems, but also creators.

In contrast to algorithmic impact assessment models, which seek to assign responsibility for the eventual harmful impacts of a certain technology to those who designed, deployed and operate it; and to privacy approaches that conceive subjects of rights as individual, autonomous agents, free and rational to make decisions, without being immersed in power dynamics that determine our ability to decide,<sup>71</sup> in a feminist approach, priority is given to the legibility of documentation (as a process and as a product) in order to build, as stated in the document *Tecnoafecciones*, “policies of co-responsibility that make us become aware, assume and collectively take charge of these interdependent relationships that are the basis of existence.”<sup>72</sup>

---

69. Margot Kaminski, “The Right to Explanation, Explained,” *Berkeley Technology Law Journal* 34, no. 1 (May 2019): 189–218, <https://doi.org/10.15779/Z38TD9N83H>

70. Emanuel Moss et al., “Assembling Accountability”.

71. Varon y Peña, “Artificial intelligence and consent”.

72. Paola Ricaurte Quijano et al., “Tecnoafecciones. Por una política de la co-responsabilidad” (Instituto Simone de Beauvoir, 2020), [https://ia601809.us.archive.org/28/items/tecnoafecciones-web/Tecnoafecciones\\_web.pdf](https://ia601809.us.archive.org/28/items/tecnoafecciones-web/Tecnoafecciones_web.pdf), 17.

## 5. Doing speculatively

Developing AI that does not reproduce oppressive logics is today only possible in the imagination, since development is only a part of the complex interplay of materials, bodies, jobs and interests involved in each piece of code that runs on a machine with learning capabilities. But as Judy Wajcman wrote years ago, “an emancipatory politics of technology requires more than hardware and software; it needs *wetware* - bodies, fluids, human agency.”<sup>73</sup>

To imagine is to open up possibilities. Sophie Toupin and Spideralex propose to talk about speculation as an active and feminist practice. *Doing speculatively*, they say,

“in a context of technology helps to prefigure the types of feminist technologies, technoscience and infrastructures needed to (re) imagine and strive for systemic transformations. *Doing speculatively* allows us to foreground feminist imaginaries that reconsider and reshape technologies, which are constituted by processes, institutions, knowledges, bodies and artefacts.”<sup>74</sup>

Imagining what we do not see of technologies allows them to affect us. In the words of Marisol Flores, “if when they tell you about smart cities instead of imagining something full of little lights and rays, as they present it to you in all the technology exhibitions, and you imagine that in reality this is a euphemism for a city full of cameras with corporate software, then this does have an impact on the decisions you make in this regard.”

Imagining that skin does not separate us but connects us, as Nadia Cortés proposes, is a way to challenge the idea that “you and I have nothing in common, [that] my individual difference dissociates us.”

“No, my difference is that I affect differently and process, respond, in different ways to the constant changes in the world. I am being the affectation with you, with all of you. There are no individuals, there

---

73. Judy Wajcman, *Technofeminism* (Madrid: Ed. Catedra, 2006), pg. 77.

74. Sophie Toupin y Spideralex, “Introduction: Radical Feminist Storytelling and Speculative Fiction: Creating new worlds by re-imagining hacking”, *ADA* 13, 2018, <https://adanewmedia.org/2018/05/issue13-toupin-spideralex>

are processes, common becomings, creations in collective to which we belong beyond our will.”<sup>75</sup>

Or, as Yásnaya Aguilar writes “place technological creation and ingenuity once again at the service of the common good, not of the market.”<sup>76</sup> Sofia Trejo says she likes “to think that we can create joint strategies at the regional level to be able to create counterweights to these large corporations because the problem is that we don't have enough computing resources now or enough resources of various things, but maybe in a collaborative way we can do that.”

For Selene Yang, “feminism in technology comes from the most basic, which would be collaborative. In other words, there is no feminist methodology that is not collaborative.” Collaborative can be understood at different levels, for example from interdisciplinary work. Referring to the problem of biases, Ximena Gutiérrez says that “a possibility is opening up to be creative in technological terms, but if you ask me, what is a solution that I see from my more idealistic or general perspective, I think that interdiscipline is a super magic key.”

She says that areas “like Natural Language Processing are more balanced in terms of gender because it is an interdisciplinary area.”

“We have written at least one all-female paper at one of the most important conferences in the area of computational linguistics and the truth is that it is not common. So my experience has been that, a working group that challenges the norm in a very technical area. I can completely notice a difference because it is a work group where beyond productivity there is a lot of empathy, understanding, etc.”

Fernanda Carles says that “on the technical side, the more social issues are seen as a waste of time. [...] People know how to use the machine but don't understand the problems.” She says that after working for a while with civil society organizations, she wanted to return to academia and do technical work, but focused on social issues, and was unable to get support. She was told “do something more technical, something that

---

75. Nadia Cortés, “Piel sintética interconectada: ficcionar para afectarnos”, at “Nos permitimos imaginar: escrituras hackfeministas para otras tecnologías”, comp. Paola Ricourte Quijano (Mexico: Instituto Simone de Beauvoir, 2020), <https://ia801805.us.archive.org/32/items/nos-permitimos-imaginar-web/NosPermitimosImaginar.pdf>, 8.

76. Aguilar Gil, “A modest proposal”.

helps the country more.” For this reason, she considers it necessary to have “more diversity in the teams, critical analysis of how these things are being used and a recognition of the limitations they also have.”

Speaking about limitations, Luciana Benotti questions that “in the area of Natural Language Processing, and in general in the areas of Artificial Intelligence, the limitations of the studies are never written, [...] researchers talk about all the good things and never about any limitations.”

“Those kinds of considerations I started to think about very recently and I am now part of the ethics committee of the Association for Computational Linguistics, to put together a bibliography of these topics and to hire specialists to come and educate us and explain, because we have no idea. There are researchers who not only have no idea, but also say, “it’s not my problem and I don’t want to understand.” I say that it is my problem and in reality this kind of spaces like the Feminist Network is what they are teaching me, how to think about these things. I think it is very important.”

And regarding diversity in teams, it is important to bring in Ruha Benjamin's critique of such solutions in the design of AI systems. She argues that “if machines are programmed to carry out tasks, both they and their designers are guided by some purpose, that is to say, intention. And in the face of discriminatory effects, if those with the power to design differently choose business as usual, then they are perpetuating a racist system”.<sup>77</sup>

In her “modest proposal to save the world”, Yásnaya Aguilar proposes a strategy she calls *tequiology*. *Tequio* (from the Nahuatl word *tequitl*) is a form of collaborative work and mutual support that has been practiced since before colonization by different peoples in the territory now known as Latin America. Also called *faena*, *kol* or *minga*, *tequio* is a fundamental social technology consisting of “small-scale, community-level labor linked into a circuit of larger tasks” and, says the author, “in the face of our current global climate emergency, we need to foster forms of technological development that emphasize living with dignity, not infinite economic growth as an end in itself. We must focus on technologies based on collaborative labor more than on competition.”<sup>78</sup>

Following this proposal, Yadira Sánchez Benítez writes about how a *tequiology* would work within the field of AI. Although because of the

---

77. Benjamin, *Race after Technology*, 60. (traducción propia)

78. Aguilar Gil, “A modest proposal to save the world”.

amount of computational resources required it is difficult to imagine a development outside of universities or corporations, she says,

“for AI infrastructures and machine learning operations to fully embrace this idea of tequiologies, collective and collaborative efforts would include not only participants involved in the design process but the entire community’s knowledge and context by following a kind of community assembly becoming more a type of collective intelligence aided by tools.”<sup>79</sup>

For a hyper-specialized field such as AI, this proposal resonates with the principles of Design Justice, especially with the fifth principle, which states, “we see the role of the designer as a facilitator rather than an expert.”<sup>80</sup> It also resonates with the commitment to co-liberation undertaken by different technology collectives in the U.S. and other territories. As Catherine D’Ignazio and Lauren Klein write,

“a model that positions co-liberation as the end goal leads to a very specific set of processes and practices, as well as criteria for success. Co-liberation is grounded in the belief that enduring and asymmetrical power relations among social groups serve as the root cause of many societal problems. Rather than framing acts of technical service as benevolence or charity, the goal of co-liberation requires that those technical workers acknowledge that they are engaged in a struggle for their own liberation as well, even and especially when they are members of dominant groups.”<sup>81</sup>

---

79. Yadira Sánchez Benítez, “A New AI Lexicon: Tequiologies”, AI Now Institute, October 22, 2021, <https://medium.com/a-new-ai-lexicon/a-new-ai-lexicon-tequiologies-38f100255820>.

80. Design Justice Network, Design Justice Network Principles, <https://designjustice.org/read-the-principles>.

81. D’Ignazio y Klein, *Data Feminism*, 141.

## 6. Towards a guide for the development of feminist AI systems

*This guide is under construction. It focuses on the preliminary stage of conception and design of an AI system, and is still very general.*

Understanding that a feminist Artificial Intelligence (AI) must respond to local needs and issues, previously identified by a group of interest that is part of the people projected as future users, this guide is not applicable to large-scale AI systems. Below we share a set of criteria, guiding questions and basic practices to keep in mind during its design.

### **Who is this basic guide oriented to?**

- Computer science teachers and students, and small companies of development and implementation of machine learning systems, committed to social transformation, and interested in integrating feminist principles in their work.
- Organizations and social movements, community initiatives and feminist collectives interested in satisfying, with an AI system, some concrete need in their work and in their territory, evaluate the relevance of such a system, and approach a group with the capacity to develop and implement it.

### **6.0. Before starting, understand AI as a result of collective processes.**

The preliminary design of an AI system is an exercise of imagination about who will use the system and how it will be used. This process is not only materialized in the product, but in the very exercise of dialogue and joint construction between the interested parties and the development team, who are committed to empathizing with the experiences, knowledge and needs of all the parties involved in the process. We summarize it in five stages.

### **6.1. Identify, through a collaborative process of knowledge exchange, what needs an organization or community wants to satisfy with an AI system.**

The objective is to reach a common understanding of the needs. For this, it is useful to have a description, as detailed as possible, of the environment where the AI system is expected to be used, and of the technical characteristics of the system. Visual representations, maps, diagrams, and even dramatizations can be used.

**Key principles:** collaboration, participation, autonomy, territorial defense, environmental justice.

**6.2. Determine, based on the recognition of the context, capabilities and available resources, whether an AI system is the best way to satisfy the identified needs.**

It is likely that an AI (understood as a deep learning system for decision making) is not the technological tool best suited to meet the needs posed by the organization or community. In that case, the development team may propose technological alternatives, or, in the process, find alternatives that do not involve digital technologies and still respond to the identified demands. The ultimate goal of this type of effort is not the deployment of certain types of technologies, but the collective construction of answers and alternatives.

**Key principles:** community development, situated knowledge, decolonial perspective, resistance.

**6.3. Agree on expectations, roles, timing and commitments at different stages of the development and deployment of an AI system.**

Collaborative work does not always involve development teams, organizations or communities fully. Defining roles, times and concrete commitments contributes to permanent and understandable communication for all parties involved, and with this, a sense of co-responsibility in the process. For example, who is in charge of coordinating communication with the development team? What about documentation? How are the beta versions of the system under development expected to be tested? Who should test it? In what spaces is the organization or community expected to become familiar with the developed tools? What other activities, besides the computational work, are involved in the development of this AI?

**Key principles:** care, alternative economies, resilience, co-responsibility.

**6.4. Recognize, within the previously agreed working group, possible risks, opportunities and conflicts associated with the development and deployment of AI systems in their different stages.**

This part involves going a little deeper into the technical characteristics of AI, and the ethical conflicts it presents, as well as the risks it could imply for the users. For example: algorithmic or datasets biases; the conditions under which existing datasets were built or the limitations involved in building one's own datasets; privacy implications for users and how to ensure continued consent in data capture and processing; licenses for use of existing codes, datasets or work environments; the costs associated with the computer equipment needed to develop and maintain a system in operation, as well as its environmental impact, among other potential conflicts.

**Key principles:** usage, intersectional view, consent.

**6.5. Define, in a conscious and informed manner by the organization or community, the resources, tools and services to be used for the development and deployment of the IA system.**

Understanding and collectively analyzing the potential conflicts associated with AI development allows us to see the advantages and disadvantages of choosing certain tools. For example, to what extent it may be beneficial or harmful to use a library under a proprietary license, or a free-to-use infrastructure, or a benchmark dataset. To make these decisions it is necessary to understand the risks inherent to the context where it will be deployed and the people who will use it, but it is also relevant to discuss the values and principles of the organization or community that will use the system.

**Key principles:** free and open source, movement building, privacy, governance.