



# What Artificial Intelligence is hiding

*Microsoft and vulnerable girls in northern Argentina*

Tomás Balmaceda, Karina Pedace and Tobias Schleider

STATE OF POWER 2023

The film *The Wizard of Oz*<sup>1</sup> premiered in 1939. One of its star actors was Terry, the dog trained to play the role of Toto, who was then considered ‘the smartest animal on the planet’. The subject of animal intelligence preoccupied many scholars of the time, while there was growing interest in understanding whether machines could think on their own. Such a possibility clearly challenged common sense, which ruled it out completely, but started to be challenged a decade after the film’s debut in the work of the British mathematician Alan Turing.<sup>2</sup> For much of the twentieth century, the idea that animals or machines were capable of thinking was considered totally preposterous—so much has changed since then!

In early 2016, the governor of Salta in Argentina chose *The Wizard of Oz* as the book to hand out to students in his province who were learning to read. The girls discovered in L. Frank Baum’s book that there is always a man behind the ‘magic’. When they became adolescents, this lesson extended to other, more concrete, areas of their lives: that it is not magic, but men that lie behind poverty, promises, disappointments and pregnancies.

By then, artificial intelligence (AI) had gone from being Turing’s challenge to become the preferred area of expertise of the world’s most powerful and influential corporations. Thanks to attractive applications on personal devices such as mobile phones and streaming platforms, it gained broad popularity.

Until a few years ago, we used to only hear the phrase ‘artificial intelligence’ to refer to HAL 9000<sup>3</sup> from *2001: A Space Odyssey* or Data, the android from *Star Trek*. But today, few are surprised by its daily use. The consensus in the media and in certain academic literature is that we are witnessing one of the most important technological revolutions in history.

However, the awe inspired by this technology—which seems to be straight out of a fairy tale or a science fiction movie—conceals its true nature: it is as much a human creation as the mechanisms that the would-be Wizard of Oz wanted to pass off as divine and supernatural events. In the hands of the state apparatus and large corporations, ‘artificial intelligence’ can be an effective instrument of control, surveillance and domination, and for consolidating the status quo. This became clear when the software giant Microsoft allied itself with the government of Salta, promising that an algorithm could be the solution to the school dropout and teenage pregnancy crisis in that region of Argentina.

## Algorithms that predict teenage pregnancy

A year after he distributed copies of *The Wizard of Oz* to the schools in his province, the governor of Salta, Juan Manuel Urtubey, announced an agreement with Microsoft’s national subsidiary to implement an AI platform designed to prevent what he described as ‘one of the most urgent problems’ in the region. He was referring to the rising number of teenage pregnancies. According to official statistics, in 2017, more than 18% of all births registered in the province were to girls under 19 years of age: 4,914 children, at a rate of more than 13 per day.

While promoting his initiative, the governor declared,<sup>4</sup> ‘We are launching a programme to prevent teenage pregnancy by using artificial intelligence with the help of a world-renowned software company. With this technology, you can predict five or six years in advance—with the first and

last name and address—which girl, future teenager has an 86% likelihood of having a teenage pregnancy’.

With almost the same fanfare as the Wizard of Oz’s greetings to visitors who found their way along the yellow brick road, Microsoft heralded the announcement<sup>5</sup> of the deal, calling it an ‘innovative initiative, unique in the country, and a major step in the province’s digital transformation process’.

A third member of the alliance between the tech giant and the government was the CONIN Foundation, headed by Abel Albino,<sup>6</sup> a medical doctor and activist who fought against the legalisation of abortion and the use of condoms.

This alliance reveals the political, economic and cultural motives behind the programme: the goal was to consolidate the concept of ‘family’ in which sex and women’s bodies are meant for reproduction—supposedly the ultimate and sacred purpose that must be protected at all cost. This well-known conservative view has been around for centuries in Latin America, but here it was dressed up in brightly coloured clothes thanks to the complicity of a US corporation Microsoft and the use of terms such as ‘artificial intelligence’ which were apparently enough to guarantee effectiveness and modernity.

The announcements also provided information on some of the methodology to be used. For example, they said that basic data ‘will be submitted voluntarily by individuals’ and allow the programme ‘to work to prevent teenage pregnancy and school dropouts. Intelligent algorithms are capable of identifying personal characteristics that tend to lead to some of these problems and warn the government’. The Coordinator of Technology of the Ministry of Early Childhood of the Province of Salta, Pablo Abeleira, declared<sup>7</sup> that ‘at the technological level, the level of precision of the model we are developing was close to 90%, according to a pilot test carried out in the city of Salta’.

What lay behind these claims?

## The myth of objective, neutral artificial intelligence

AI has already become embedded in not only public discourse but also our daily lives. It sometimes seems as if everyone knows what we mean by ‘artificial intelligence’ (AI). However, this term is by no means unambiguous, not only because it is usually used as an umbrella under which very similar and related—but not synonymous—concepts appear—such as ‘machine learning’, ‘deep learning’ or cognitive computing, among others—but also because a closer analysis reveals that the very concept of intelligence in this context is controversial.

In this essay, we will use AI to refer to algorithm models or systems that can process large volumes of information and data while ‘learning’ and improving their ability to perform their task beyond what they were originally programmed to do. A case of AI, for instance, is an algorithm that after processing hundreds of thousands of pictures of cats can extract what it needs to recognise a cat in a new photo, without mistaking it for a toy or cushion. The more photographs it is fed, the more it will learn and the fewer mistakes it will make.

These developments in AI are sweeping the globe and are already used in everyday technology such as voice recognition of digital assistants like Siri and Alexa, as well as in more ambitious projects including driverless cars or tests for the early detection of cancer and other diseases. There is a very wide range of uses for these innovations, which affects many industries and sectors of society. In the economy, for example, algorithms promise to identify the best investments on the stock market. In the political arena, there have been social media campaigns for or against an electoral candidate that have been designed to appeal to different individuals based on their preferences and internet use. In relation to culture, streaming platforms use algorithms to offer personalised recommendations on series, movies or music.

The success of these uses of the technology and the promises of benefits that, until a few years ago, existed only in science fiction, has inflated the perception of what AI is really capable of doing. Today, it is widely regarded as the epitome of rational activity, free of bias, passions and human error.

This is, however, just a myth. There is no such thing as 'objective AI' or AI that is untainted by human values. Our human—perhaps all too human—condition will inevitably have an impact on technology.

One way to make this clear is to remove some of the veils that conceal a term such as 'algorithm'. The philosophy of technology allows us to distinguish at least two ways of defining it in conceptual terms. In a narrow sense, an algorithm is a *mathematical construct* that is selected because of its previous effectiveness in solving problems similar to those to be solved now (such as deep neural networks<sup>8</sup>, Bayesian networks<sup>9</sup> or 'Markov chains'<sup>10</sup>). In the *broad* sense, an algorithm is a *whole technological system* comprising several inputs, such as training data, which produces a statistical model designed, assembled and implemented to resolve a pre-defined practical issue.

It all begins with a simplistic understanding of data. Data emerge from a process of selection and abstraction and consequently can never offer an objective description of the world. Data are inevitably partial and biased since they are the result of human decisions and choices, such as including certain attributes and excluding others. The same happens with the notion of data-based forecasting. A key issue for government use of data-based science in general and machine learning in particular is to decide what to measure and how to measure it based on a definition of the problem to be addressed, which leads to choosing the algorithm, in the narrow sense, that is deemed most efficient for the task, no matter how deadly the consequences.<sup>11</sup> Human input is thus crucial in determining which problem to solve.

It is thus clear that there is an inextricable link between AI and a series of human decisions. While machine learning offers the advantage of processing a large volume of data rapidly and the ability to identify patterns in the data, there are many situations where human supervision is not only possible, but necessary.

## Pulling back the curtain on AI

When Dorothy, the Tin Man, the Lion and the Scarecrow finally met the Wizard of Oz, they were fascinated by the deep, supernatural voice of this being who, in the 1939 film version, was played by Frank Morgan and appeared on an altar behind a mysterious fire and smoke. However, Toto, Dorothy's dog, was not as impressed and pulled back the curtain, exposing the sham: there was someone manipulating a set of levers and buttons and running everything on stage. Frightened and embarrassed, the would-be wizard tried to keep up the charade: 'Pay no attention to the man behind the curtain!' But, when cornered by the other characters, he was forced to admit that it was all a hoax. 'I am just a common man', he confessed to Dorothy and her friends. The Scarecrow, however, corrected him immediately: 'You're more than that. You're a humbug'.

When we take away the fancy clothes and gowns, we see AI for what it really is: a product of human action that bears the marks of its creators. Sometimes, its processes are seen as being similar to human thought, but are treated as devoid of errors or bias. In the face of widespread, persuasive rhetoric about its value-neutrality and the objectivity that goes along with it, we must analyse the inevitable influence of human interests at various stages of this supposedly 'magic' technology.

Microsoft and the government of Salta's promise to predict 'five or six years in advance, with names, last names and addresses, which girl or future adolescent has an 86% likelihood of having a teenage pregnancy' ended up being just an empty promise.

The fiasco began with the data: they used a database collected by the provincial government and civil society organisations (CSOs) in low-income neighbourhoods in the provincial capital in 2016 and 2017. The survey reached just under 300,000 people, of whom 12,692 were girls and adolescents between 10 and 19 years of age. In the case of minors, information was gathered after obtaining the consent of 'the head of family' (sic).

These data were fed into a machine-learning model that, according to its implementers, is able to predict with increasing accuracy which girls and adolescents will become pregnant in the future. This is absolute nonsense: Microsoft was selling a system that promised something that is technically impossible to achieve.<sup>12</sup> It was fed a list of adolescents who had been assigned a likelihood of pregnancy. Far from enacting any policies, the algorithms provided information to the Ministry of Early Childhood so it could deal with the identified cases.

The government of Salta did not specify what its approach would entail, nor the protocols used, the follow-up activities planned, the impact of the applied measures—if indeed the impact had been measured in some way—the selection criteria for the non-government organisations (NGOs) or foundations involved, nor the role of the Catholic Church.

The project also had major technical flaws: an investigation by the World Web Foundation<sup>13</sup> reported that there was no information available on the databases used, the assumptions underpinning the design of the models, or on the final models were designed, revealing the opacity of the process. Furthermore, it affirmed that the initiative failed to assess the potential inequalities and did not pay special attention to minority or vulnerable groups that could be affected. It also did not consider the difficulties of working with such a wide age group in the survey and the risk of discrimination or even criminalisation.

The experts agreed that the assessment's data had been slightly contaminated, since the data used to evaluate the system were the same ones used to train it. In addition, the data were not fit for the stated purpose. They were taken from a survey of adolescents residing in the province of Salta that requested personal information (age, ethnicity, country of origin, etc.) and data on their environment (if they had hot water at home, how many people they lived with, etc.) and if they had already been or were currently pregnant. Yet, the question that they were trying to answer based on this *current* information was whether a teenage girl might get pregnant *in the future*—something that seemed more like a premonition than a prediction. Moreover, the information was biased, because data on teenage pregnancy tend to be incomplete or concealed given the inherently sensitive nature of this kind of issue.

Researchers from the Applied Artificial Intelligence Laboratory of the Computer Sciences Institute at the University of Buenos Aires found that in addition to the use of unreliable data, there were serious methodological errors in Microsoft's initiative. Moreover, they also warned of the risk of policymakers adopting the wrong measures: 'Artificial intelligence techniques are powerful and require those who use them to act responsibly. They are just one more tool, which should be complemented by others, and in no way replace the knowledge or intelligence of an expert', especially in an area as sensitive as public health and vulnerable sectors.<sup>14</sup>

And this raises the most serious issue at the centre of the conflict: even if it were possible *to predict* teenage pregnancy (which seems unlikely), it is not clear *what purpose* this would serve. *Prevention* is lacking throughout the entire process. What it did do, however, is create an inevitably high risk of stigmatising girls and adolescents.

## AI as an instrument of power over vulnerable populations

From the outset, the alliance between Microsoft, the government of Salta and the CONIN Foundation was founded on preconceived assumptions that are not only questionable, but also in conflict with principles and standards enshrined in the Argentinean Constitution and the international conventions incorporated into the national system. It is unquestionably based on the idea that (child or teenage) pregnancy is a disaster, and in some cases the only way to prevent it is through direct interventions. This premise is linked to a very vague stance on the attribution of responsibility.

On the one hand, those who planned and developed the system appear to see pregnancy as something for which no one is responsible. Yet, on the other hand, they place the responsibility exclusively on the pregnant girls and adolescents. Either way, this ambiguity contributes, first of all, to the objectification of the people involved and also renders invisible those who are in fact responsible: primarily the men (or teenagers or boys, but mainly men) who obviously contributed to the pregnancy (people often say, with a crude, euphemistic twist, that the girl or teenager 'got herself pregnant'). Second, it ignores the fact that in most cases of pregnancy among young women and in *all* cases of pregnancy among girls, not only is it wrong to presume that the girl or teenager consented to sexual intercourse, but this assumption should be completely ruled out. In sum, this ambiguous stance obscures the crucial fact that all pregnancies of girls and many pregnancies of young women are the result of rape.

With regard to the most neglected aspect of the system—that is, predicting the school dropout rate—it is assumed (and concluded) that a pregnancy will inevitably lead a pupil to drop out of school. While the opportunity cost that early pregnancy and motherhood imposes on women should never be ignored, the interruption or abandonment of formal education is not inevitable. There are examples of inclusive programmes and policies that have been effective in helping to avoid or reduce dropout rates.

From a broader perspective, the system and its uses affect rights that fall within a spectrum of sexual and reproductive rights, which are considered human rights. Sexuality is a central part of human development, irrespective of whether individuals choose to have children. In the case of minors, it is important to take account of differences in their evolving capacities, while bearing in mind that the guidance of their parents or guardians should always give priority to their capacity to exercise rights on their own behalf and for their own benefit. Sexual rights in particular entail specific considerations. For example, it is essential to respect the particular circumstances of each girl, boy or adolescent, their level of understanding and maturity, physical and mental health, relationship with various family members and ultimately the immediate situation they are facing.

The use of AI has concrete impacts on the rights of (potentially) pregnant girls and teenagers. First, the girls and teenagers the right to personal autonomy was violated. We have already mentioned their objectification and the project's indifference towards their individual interests in pursuit of a supposed general interest. The girls and teenagers were not even considered as rights holders and their individual desires or preferences were completely ignored.

In this Microsoft project, AI was used as an instrument to yield power over girls and teenage women, who were catalogued without their consent (or their knowledge, apparently). According to those who promoted the system, the interviews were held with the 'heads of the family' (especially their fathers) without even inviting them to participate. Moreover, the questionnaires included highly personal matters (their intimacy, sex lives, etc.) on which their parents would seldom be able to respond in detail without invading their daughter's privacy or—just as serious—relying on assumptions or biases that the state would then assume to be true and legitimate.

Other violations include the rights to intimacy, privacy and freedom of expression or opinion, while the rights to health and education are at risk of being ignored, despite the declarations of the authorities and Microsoft about their intention to take care of the girls and adolescent women. Finally, it is worth mentioning a related right that is of particular importance in the specific context of this project: the right to freedom of thought, conscience and religion.

We would not go so far as to claim that this episode had a happy ending, like the Wizard of Oz did. But the Microsoft project did not last long. Its interruption was not because of criticisms from activists, however, but for a much more mundane reason: in 2019, national and state elections were held in Argentina and Urtubey was not re-elected. The new administration terminated several programmes, including the use of algorithms to predict pregnancy, and reduced the Ministry of Early Childhood, Childhood and Family to the status of a secretariat.

## What AI is hiding

The rhetorical smoke and mirrors of value-neutral and objective AI developments fall apart when challenged by voices that assert that this is impossible in principle, as we argued in the first section, given the participation of human analysts in several stages of the algorithms' development. Men and women defined the problem to be resolved, designed and prepared the data, determined which machine-learning algorithms were the most appropriate, critically interpreted the results of the analysis and planned the proper action to take based on the insights that the analysis revealed.

There is insufficient reflection and open discussion about the undesirable effects of the advance of this technology. What seems to prevail in society is the idea that the use of algorithms in different areas not only guarantees efficiency and speed, but also the non-interference of human prejudices that can 'taint' the pristine action of the codes underpinning the algorithms. As a result, people take for granted that AI has been created to improve society as a whole or, at least certain processes and products. But hardly anyone questions the basics—for whom will this be an improvement, who will benefit and who will assess the improvements? Citizens? The state? Corporations? Adolescent girls from Salta? The adult men who abused them? Instead, there is a lack of real awareness about the scale of its social impact or the need to discuss whether such a change is inevitable.

People are no longer surprised by the constant news on the introduction of AI into new fields, except for what is new about it, and just like the passage of time, it is treated as something that cannot be stopped or revisited. The growing automation of the processes that humans used to carry out may generate alarm and concern, but it does not spark interest in halting it or reflecting on what the future of work and society will be like once AI takes over much of our work. This raises a number of questions that are rarely asked: Is this really desirable? For which social sectors? Who will benefit from greater automation and who will lose out? What can we expect from a future where most traditional jobs will be performed by machines? There seems to be neither the time nor the space for discussing this matter: automation simply happens and all we can do is to complain about the world we have lost or be amazed by what it can achieve today.

This complacency with the constant advances of technology in our private, public, work and civic lives is thanks to trust in the belief that these developments are 'superior' to what can be achieved through mere human effort. Accordingly, since AI is much more powerful, it is 'smart' (the 'smart' label is used for mobile phones, vacuum cleaners and coffeemakers, among other objects that would make Turing blush) and free of biases and intentions. However, as pointed out earlier, the very idea of value-neutral AI is a fiction. To put it simply and clearly: there are biases in all stages of algorithm design, testing and application and it is therefore very difficult to identify them and even harder to correct them. Nonetheless, it is essential to do so in order to unmask its supposedly sterile nature devoid of human values and errors.

An approach focused on the dangers of AI, along with an optimistic stance about its potential, could lead to an excessive dependence on AI as a solution to our ethical concerns—an approach where AI is asked to answer the problems that AI had produced. If problems are considered as purely technological, they should only require technological solutions. Instead we have human decisions dressed up in technological dress. We need a different approach.



The case of the algorithms that were supposed to predict teenage pregnancies in Salta exposes how unrealistic the image of the so-called objectivity and neutrality of artificial intelligence is. Like Toto, we cannot ignore the man behind the curtain: the development of algorithms is not neutral, but rather based on one decision made from many possible choices. Since the design and functionality of an algorithm reflects the values of its designers and its intended uses, algorithms inevitably lead to biased decisions. Human decisions are involved in defining the problem, data preparation and design, the selection of the type of algorithm, the interpretation of the results and the planning of actions based on their analysis. Without qualified and active human supervision, no AI algorithm project is able to achieve its objectives and be successful. Data science works best when human experience and the potential of algorithms work in tandem.

Artificial intelligence algorithms are not magic, but they do not need to be a hoax, as the Scarecrow argued. We just have to recognise that they are human.

*Translated by Karen Lang*

## **BIOGRAPHIES**

---

Karina Pedace teaches graduate and post-graduate students at the University of Buenos Aires and National University of Matanza, and is a researcher at the Institute of Philosophical Research of Argentinean Society (IIF-SADAF- CONICET). Her current research areas include the philosophy of technology, metaphysics of the mind and research methodologies. She is executive secretary of the UNESCO Latin American Network of Women Philosophers and co-founder of the Artificial Intelligence, Philosophy and Technology Research Group (GIFT). In 2022, she was recognised internationally as one of 100 Brilliant Women in AI Ethics <https://womeninaethics.org/the-list/of-2022/>

Tomás Balmaceda has a PhD in Philosophy from the University of Buenos Aires. Currently he is Researcher in IIF (SADAF/CONICET) and part of the GIFT group which analyses technology and artificial intelligence through the lens of philosophy. Author of various books, his interests include the ethics of network influence, new longevity and financial education for the LGBTIQ+ population.

Tobías J. Schleider is a lawyer and specialist in criminal law for the National University of Mar de Plata and Doctor of the University of Buenos Aires in Philosophy of Law. He is a Professor at the National University of the South, where he directs the degree in Public Security. His current lines of research include technology-supported violence prevention, human action theory, causality and the influence of luck in the attribution of responsibility.

# Endnotes

- 1 Internet Archive (2021) Available at: <https://archive.org/details/the-wizard-of-oz-1080p>. (Accessed 13 March 2023).
- 2 'Alan Turing' (2023) *Wikipedia, the Free Encyclopaedia*. Available at: [https://en.wikipedia.org/wiki/Alan\\_Turing](https://en.wikipedia.org/wiki/Alan_Turing) (Accessed 13 March 2023).
- 3 'Hal 9000' (2023) *Wikipedia, the Free Encyclopaedia*. Available at: [https://en.wikipedia.org/wiki/HAL\\_9000](https://en.wikipedia.org/wiki/HAL_9000) (Accessed 13 March 2023).
- 4 Página12 (2018) *La inteligencia artificial de Urtubey*. Available at: <https://www.pagina12.com.ar/107412-la-inteligencia-artificial-de-urtube> (Accessed 13 March 2023)
- 5 Microsoft News Center LATAM (2017) Microsoft y el gobierno de Salta firman un acuerdo para aplicar la inteligencia artificial en la prevención de los problemas más urgentes. Available at: <https://news.microsoft.com/es-xl/microsoft-gobierno-salta-firman-acuerdo-aplicar-la-inteligencia-artificial-la-prevencion-los-problemas-mas-urgentes/> (Accessed 13 March 2023)
- 6 Perfil (2018) Albino dijo que el preservativo no protege porque "el virus del SIDA atraviesa la porcelana" Available at: <https://www.perfil.com/noticias/politica/albino-dijo-que-el-preservativo-no-protege-del-vih-porque-atravesia-la-porcelana.phtml> (Accessed 13 March 2023)
- 7 Alonso, A. (2018) Microsoft democratiza la IA y los servicios cognitivos, It.sitio, 28 March. Available at: <https://www.itsitio.com/es/microsoft-democratiza-la-ia-y-los-servicios-cognitivos/> (Accessed 13 March 2023)
- 8 'Deep neural networks' (2023) *Wikipedia, the Free Encyclopaedia*. Available at: [https://en.wikipedia.org/wiki/Deep\\_learning#Deep\\_neural\\_networks](https://en.wikipedia.org/wiki/Deep_learning#Deep_neural_networks) (Accessed 13 March 2023).
- 9 'Bayesian networks' (2023) *Wikipedia, the Free Encyclopaedia*. Available at: [https://en.wikipedia.org/wiki/Bayesian\\_networks](https://en.wikipedia.org/wiki/Bayesian_networks) (Accessed 13 March 2023).
- 10 'Markov chains' (2023) *Wikipedia, the Free Encyclopaedia*. Available at: [https://en.wikipedia.org/wiki/Markov\\_chains](https://en.wikipedia.org/wiki/Markov_chains) (Accessed 13 March 2023).
- 11 Brunet, P., Font, T., Rodríguez, J. (2022) Robots asesinos: 18 preguntas y respuestas. Available at: <https://centredelas.org/publicacions/robots-asesinos-18-preguntas-y-respuestas/?lang=es> (Downloaded 13 March 2023).
- 12 Eubanks, V. (2018) *Automating Inequality; How high-tech tools profile, police and punish the poor*. New York: St Martin's Press.
- 13 Ortiz Freuler, J. and Iglesias, C. (2018) *Algoritmos e Inteligencia Artificial en Latinoamérica: Un Estudio de implementaciones por parte de Gobiernos en Argentina y Uruguay*. World Wide Web Foundation. Available at: [https://webfoundation.org/docs/2018/09/WF\\_AI-in-LA\\_Report\\_Spanish\\_Screen\\_AW.pdf](https://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Spanish_Screen_AW.pdf) (Accessed 13 March 2023)
- 14 Laboratorio de Inteligencia Artificial Aplicada (2018) *Sobre la predicción automática de embarazos adolescentes*. Available at: <https://liaa.dc.uba.ar/es/sobre-la-prediccion-automatizada-de-embarazos-adolescentes/>