

GLOBAL INFORMATION SOCIETY WATCH 2019

Artificial intelligence: Human rights, social justice and development



ASSOCIATION FOR PROGRESSIVE COMMUNICATIONS (APC),
ARTICLE 19, AND SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY (SIDA)

Global Information Society Watch

2019



Global Information Society Watch 2019

Artificial intelligence: Human rights, social justice and development

Operational team

Valeria Betancourt (APC)
Alan Finlay (APC)
Mallory Knodel (ARTICLE 19)
Vidushi Marda (ARTICLE 19)
Maja Romano (APC)

Project coordination team

Valeria Betancourt (APC)
Cathy Chen (APC)
Flavia Fascendini (APC)
Alan Finlay (APC)
Mallory Knodel (ARTICLE 19)
Vidushi Marda (ARTICLE 19)
Leila Nachawati (APC)
Lori Nordstrom (APC)
Maja Romano (APC)

GISWatch 2019 advisory committee

Namita Aavriti (APC)
Rasha Abdul Rahim (Amnesty International)
Alex Comminos (Research ICT Africa)
Malavika Jayaram (Digital Asia Hub)
J. Carlos Lara (Derechos Digitales - América Latina)
Joy Liddicoat (Centre for Law and Emerging Technologies, University of Otago)
Andrew Lowenthal (EngageMedia)
Micaela Mantegna (Geekylegal/Machine Intelligence Lab, Center for Technology and Society, San Andres University)
Valeria Milanés (Asociación por los Derechos Civiles)

Project coordinator

Maja Romano (APC)

Editor

Alan Finlay (APC)

Assistant editor and proofreading

Lori Nordstrom (APC)

Publication production support

Cathy Chen (APC)

Graphic design

Monocromo

Cover illustration

Matías Bervejillo

We would like to extend a special note of thanks to a number of authors who have made ad honorem contributions to this edition of GISWatch.

We gratefully acknowledge the following:

Philip Dawson and Grace Abuhamad (Element AI)
Anita Gurumurthy and Nandini Chami (IT for Change)
Rasha Abdul Rahim (Amnesty International)



APC would like to thank the Swedish International Development Cooperation Agency (Sida) and ARTICLE 19 for their support for Global Information Society Watch 2019.

Published by APC

2019

Printed in USA

Creative Commons Attribution 4.0 International (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0/>

Some rights reserved.

Global Information Society Watch 2019 web and e-book

ISBN 978-92-95113-13-8

APC Serial: APC-201910-CIPP-R-EN-DIGITAL-302

Disclaimer: The views expressed herein do not necessarily represent those of Sida, ARTICLE 19, APC or its members.

LATIN AMERICA

VICTOR FRANKENSTEIN'S RESPONSIBILITY? DETERMINING AI LEGAL LIABILITY IN LATIN AMERICA



Franco Giandana (with David Morar)

Creative Commons Argentina; Universidad Nacional de Córdoba
www.unc.edu.ar

Introduction

The story of Victor Frankenstein, the scientist who lost control of his creation, is a great starting point to ask: how do we make artificial intelligence (AI) developers responsible for the software they create, and for any subsequent potential harm it causes? Was Frankenstein guilty for the harm his creation caused? What happens when it is impossible to control the AI systems that are being implemented? Do we need to revise our frameworks? Is civic responsibility in AI objective or subjective? How do we determine legal causation?

These questions have become particularly relevant in Latin America, where a new AI system developed by the Artificial Intelligence Laboratory of the University of Buenos Aires has been implemented in the judicial system in Buenos Aires, the Constitutional Court in Colombia, and at the Inter-American Court of Human Rights in Costa Rica.

The system, called Prometea,¹ is helping authorities resolve “simple” cases in different fields, for example, cases related to the right to housing, to individuals in vulnerable conditions or with disabilities, involving labour rights or children and adolescents, to do with road safety, or price control for public contracts. Even criminal cases are dealt with by the system.

While civil society has raised concerns over how this project will guarantee due process, many other cities in the region have already expressed interest in deploying similar narrow AI systems.²

In the context of the Prometea system, this report broadens the discussion to outline some of the legal considerations that the courts face when dealing with AI-related harm.

The rule of law or the rule of AI?

Equality before the law is not a reality in Latin America. There are still strong, concentrated elites that hold the power, both in public office and in the private sector. AI could, potentially, help in closing the space between those in power and the populace who seek, through public institutions, the fulfilment of their rights.

Prometea is operated using voice recognition, assisting with different tasks, from providing possible solutions to cases to managing procedures and processes. The operator speaks to the system to give it instructions, which the software processes at a very high speed, reducing, as the Inter-American Court of Human Rights has shown, a workload of three days to two minutes.

Prometea uses a machine-learning prediction system. It first analyses thousands of documents that have been organised into categories in order to help the system “learn” from them. From there any new file that is entered into the system can be evaluated by Prometea, which determines where strong case history exists, offering a “ruling” on the new case, which is then approved by a judge. The system has a 96% “success rate” in that its rulings are accepted nearly all the time. As stated by the Prometea developers in an article published on the Bloomberg Businessweek website: “Prometea is being used for stuff like taxi license disputes, not murder trials, but it’s a significant automation of the city’s justice system.”³ From 151 judgments signed using Prometea at the Attorney General’s Office in Buenos Aires, 97 found the solution solely by using this system, while 54 cases only used it as a virtual assistant to automate tasks related to the legal procedures involved.

But equality before the law is not a mathematical expression. It is a more complex, emotional and holistic perspective for judging everyone according to the same behavioural standards (i.e. the law), allowing them to be able to access the justice system, and to seek recourse when the law is infringed.

¹ See also the country report from Colombia in this edition of GISWatch.

² “Narrow AI refers to AI which is able to handle just one particular task. A spam filtering tool, or a recommended playlist from Spotify, or even a self-driving car – all of which are sophisticated uses of technology – can only be defined via the term ‘narrow AI’.” Trask. (2018, 2 June). General vs Narrow AI. *Hackernoon*. <https://hackernoon.com/general-vs-narrow-ai-3d0do2ef3e28>

³ Gillespie, P. (2018, 26 October). This AI Startup Generates Legal Papers Without Lawyers, and Suggests a Ruling. *Bloomberg Businessweek*. <https://www.bloomberg.com/news/articles/2018-10-26/this-ai-startup-generates-legal-papers-without-lawyers-and-suggests-a-ruling>

Law – that social instrument we humans have built to ensure that people can co-exist in harmony – must always contain visible, human-embedded constructions of the world. This is a core element of the rule of law; it makes it possible, and it allows legal liability and responsibility to be enforced by the state. In this way, the institutions of justice have the legitimacy they need to be trusted and valued by our societies.

As is the case, though, with anything laced with humanity, flaws abound. And they do so in legislative processes as well, at every step of the way during due process. Because of this, one can easily argue that justice, inherently flawed, is an evolving concept that needs to be strictly monitored and improved. This comes with the important caveat that these improvements must not be blind and simply for the sake of making “improvements”.

However, a deeper look at AI would show that not only are we removing the visible human perspective from the equation, we are further embedding an invisible, unattributed and unspoken human element. The AI creations that would compute cases and carry out analysis are not simply ones and zeros. They are subject to the thoughts, whims and biases of their creators. The use of AI therefore doubly weakens the position of the judge, and as a consequence, the rule of law.

How sufficient is the current legal framework to deal with AI?

Justice has always been a value decisive to any human organisation, and there have been different ways to deliver justice throughout history. The concept of justice is somewhat flexible and evolving, so bringing in modifications to deliver a better quality of justice is not something unimaginable. However, these innovations deserve a detailed analysis in order to protect basic human rights. But before we analyse how implementing AI to assist judicial institutions might impact our concept of the rule of law, let us briefly analyse how prepared our legal national frameworks are for facing the challenge of AI technologies in general.

Legal liability for damages suffered due to AI systems is clearly a challenge to traditional law, especially in regions where judges and policy makers are lacking the sufficient knowledge to comprehend the potential and relevance of AI. Nevertheless, before we ask for new regulations and better legislation, we must first review the actual state of our legal frameworks and how resilient they are to the implementation of new technologies.

Weaknesses in data protection laws

Data protection laws already frame what rights must be protected when AI systems collect and

process data. In countries like Brazil and Argentina, general data protection laws state a clear difference between personal data and sensitive personal data, determining boundaries and limitations for the collection and usage for every kind of data. But even when there are clear standards to determine whether certain data is to be considered sensitive or not, AI is capable of inferring and generating sensitive information from non-sensitive data. At the same time, the law specifies the importance of the data to be collected complying with the principle of consent. However, there are broad exceptions to the need for consent for certain objectives, for example, for national security purposes (e.g. article 4.III.a of the Brazilian General Data Protection Law), leaving a broad scope for interpretation that often escapes public debate and can be potentially used for implementing risky technologies.

A problem with the right to explanation

Only a specific, well-constructed right of explanation standard can allow AI to properly assist judges in the process of ruling on legal matters. So far, however, we also are not aware of how a potential right of explanation would work, since a legal explanation is enough to legitimise judicial resolutions. Citizens have the right to understand a judge’s decision, and legal explanation is no longer enough when AI systems are involved. The transparency of the systems implemented in the public sphere is of the utmost importance to make understanding and legitimacy possible. It is important that these systems are permanently monitored and tested by independent committees that understand both the technical and legal features that legitimise their use. If we let algorithms determine what kind of consequences our actions legally have, we will be moving away from the standard of the rule of law and entering a judicial terrain that is mostly assisted, and decided on, by algorithms.

The problem of proportionate information

One problem with using AI technology in the judicial system is that there is no standard on how much information is adequate and necessary considering the purpose it is being collected for, even for simple repetitive cases such as those Prometea is used for, simply because human behaviour is complex and can express itself in multiple different scenarios. Any limitation of the information needed to justify or defend the subjects involved will result in arbitrary judgment (of course, collecting too much data can cause worse scenarios). It also creates the illusion that every document that Prometea uses for training is rightful without a competent analysis. There is a

high possibility that out of 2,400 judgements used as a training set at the Attorney General's Office in Buenos Aires, a portion of them received some level of unfair treatment in their ruling and therefore there is a logical concern for a detailed analysis on every single document provided to train Prometea.

The legal responsibility of the AI developer

Given that AI constantly changes the algorithms in ways the developer cannot always foresee, how can they be responsible if their AI system causes damages?

The responsibility of an AI developer can be determined by two factors. One is the creation of a risk which is irrespective of any intention of those who designed, developed and implement the system. This is what we know as objective responsibility. Here it is always important to determine if the system is capable of conducting illegal actions *ab initio*, that is, from the outset of the development. For example, a developer would need to run regular risk assessments to identify, eliminate or reduce the possibilities of a system operating through bias or using information in a discriminatory way. Creating a creature like Dr. Frankenstein's could easily represent a risk, and if the "monster" suddenly turns uncontrollable, the scientist would be liable solely because of creating the risk.

The second factor focuses on the positive obligation that everyone has to not harm or violate anyone's rights. So if an AI system causes damages, the developers will be liable only if they have acted unlawfully because they have been negligent at some point of the AI life cycle, or because they had the intention to cause harm. In this case, the scientist behind Frankenstein's monster would not be responsible if he operated diligently.

Determining how responsibility will be attributed to those liable depends on what kind of system has produced the harm, the nature of the activity it was assisting, and how the harm was produced. With regard to objective responsibility, we understand it could be applied only to those systems that execute an activity which is already potentially risky. For the AI system that does not execute these kinds of activities, a subjective attribution should be considered, introducing the need to establish the nexus between the harmful outcome and the developer's actions, even when the oversight power is reduced as the system runs and the algorithms change.

Finally, since most people in Latin America do not have the opportunity to sue companies based in the European Union or in North America under their jurisdictions, as these are expensive legal

procedures, many will have to choose their own jurisdictions, even with the delays, the random inefficiencies, and the lack of knowledge on how AI systems truly work.

Conclusion

New technologies such as AI, while offering an opportunity for innovation in our societies, including our legal systems, are raising a number of critical questions with respect to their application. Flawed digital technologies are increasingly at the core of our daily activities, and they interact with us. These technologies act like a substitute intelligence to which we can delegate tasks, ask for directions or answers to complex problems, unfolding the nature of reality by analysing data in ways we, as humans, cannot achieve alone.

Now, as we have seen, AI is already being used in our judicial system. Prometea is a first step towards the implementation of AI in the judicial system, and even if its only scope is "simple" cases, we are aware that this scope might widen if it keeps functioning "efficiently".

This success might lead to broader use of the technology in more complex stages of the judicial process, something the developers of the system already say they are looking forward to happening in the near future.⁴

Latin American countries face the challenge of fully integrating the everyday activities of their citizens with these new digital paradigms, while also protecting the unique characteristics and needs of a vast, diverse group of people in the region. After two centuries of somewhat stable rule of law, we have created a strong notion of the importance of our institutions and laws to guarantee the exercise of human rights. Even if AI offers a way of bringing the judiciary closer to the people, it needs to be implemented in a way that safeguards our shared sense of the importance of institutions and the law – this sense of importance is fundamental to our trust in these institutions.

Even when our current legal frameworks need revision and possible modification, we already have the legal means to protect our rights, with institutions and processes that offer legal recourse. Personal data, our privacy and freedom of expression find protection in our legal frameworks. The institutions of *habeas data*,⁵ constitutional control

4 Murgo, E. (2019, 17 May). Prometea, Inteligencia Artificial para agilizar la justicia. *Unidiversidad*. www.unidiversidad.com.ar/prometea-inteligencia-artificial-para-agilizar-la-justicia

5 Habeas data is a constitutional remedy to rectify, protect, actualise or erase the data and information of an individual, collected by public or private subjects using manual or automated methods.

or the ordinary mechanisms to seek recourse for the harm caused by AI are already here.

But simply having the right laws at the right time is not necessarily always a winning formula. A strong, competent judicial system that understands what is at stake and how to respond to a highly technological age is needed. Procedural law and constitutional law must insure that no changes are made to due process without transparency. As Jason Tashea wrote for *Wired* magazine:

How does a judge weigh the validity of a risk-assessment tool if she cannot understand its decision-making process? How could an appeals court know if the tool decided that socioeconomic factors, a constitutionally dubious input, determined a defendant's risk to society?⁶

Transparency is needed to protect due process and ultimately, the rule of law. If it is not provided – for instance, if the decision-making codes behind algorithms are protected as industrial secrets or intellectual property – due process and the rule of law will be in danger. Because of this, a broad understanding of the legal, ethical and rights repercussions of such a deployment should be sought.

Action steps

The following action steps are suggested for Latin America, and elsewhere in the world:

- Build creative narratives to communicate the risks and the repercussions of implementing AI systems within the public sphere.
- Design a capacity-building agenda for citizens to strengthen their right to due process in courts using Prometea or any similar system.
- Seek collaboration with other organisations and specialists in the region to build general consensus on the ethical use of Prometea and provide a broader understanding of the challenges it represents.
- Advocate for policy reform in order to include specific regulations on how AI should protect rights and how transparency should be realised in AI systems.
- Foster strategic litigation against AI violating human rights. Present a detailed set of evidence to support your claim.

⁶ Tashea, J. (2017, 17 April). Courts are using AI to sentence criminals. That must stop now. *Wired*. <https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now>

Artificial intelligence: Human rights, social justice and development

Artificial intelligence (AI) is now receiving unprecedented global attention as it finds widespread practical application in multiple spheres of activity. But what are the human rights, social justice and development implications of AI when used in areas such as health, education and social services, or in building “smart cities”? How does algorithmic decision making impact on marginalised people and the poor?

This edition of Global Information Society Watch (GISWatch) provides a perspective from the global South on the application of AI to our everyday lives. It includes 40 country reports from countries as diverse as Benin, Argentina, India, Russia and Ukraine, as well as three regional reports. These are framed by eight thematic reports dealing with topics such as data governance, food sovereignty, AI in the workplace, and so-called “killer robots”.

While pointing to the positive use of AI to enable rights in ways that were not easily possible before, this edition of GISWatch highlights the real threats that we need to pay attention to if we are going to build an AI-embedded future that enables human dignity.

GLOBAL INFORMATION SOCIETY WATCH
2019 Report
www.GISWatch.org

