digital impact alliance

spotlight

Algorithms in Government

A Magic Formula or a Divisive Force?

Prasanna Lal Das
November 2022

Key Insights

1. **Algorithmic decision-making is becoming prevalent in the public sector worldwide,** and governments in developing countries are increasingly beginning to deploy algorithms to deliver citizen and business services as part of their digital transformation agenda.

2. **Many algorithmic decision-making initiatives in developing countries are still at an early stage,** as the case studies in this issue brief suggest. The examples featured in this brief are local and carefully designed, with data governance challenges such as privacy and data security in mind.

3. **Developing countries face several distinct data governance challenges** related to the design and implementation of algorithmic decision-making services.

    a. Institutions in developing countries have an **extreme legitimacy, accountability, and transparency problem.**
    b. Poor local data means that **people in developing countries are inadequately represented in training data.**
    c. People in developing countries have **less experience in interacting with machines and algorithms, and there's a scarcity of data in local languages** to close the cultural gap.
    d. Developing countries have had **limited involvement in developing standards for fairness, transparency, and accountability** in algorithmic decision-making.
    e. Developing countries are **dependent on international data infrastructure** to develop and manage their algorithms.
    f. Developing countries deploying algorithmic decision-making are **dependent on big tech companies** but have little leverage over them.

4. **Opportunities to address these specific data governance challenges are emerging, including:**

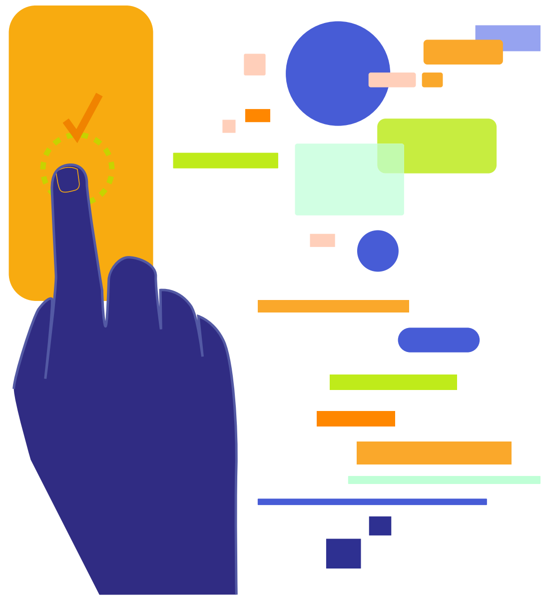
    a. Create **regional or other data alliances** to tackle relevant data governance challenges.
    b. Focus on **cases that don't depend on personal data** to deliver relevant services to citizens and businesses.
    c. Keep the **emphasis on people,** both as designers and supervisors of algorithms and as consumers of algorithmic services.

5. **Many additional data governance challenges posed by algorithmic decision-making can be addressed as part of a country's overall digital transformation agenda.** These are not the focus of this issue brief but include themes such as the overall legal/regulatory/enabling environment, infrastructure development, financing, capacity/skills development, and institutional support.

SECTION 1

Introduction

It's easy to be seduced by the power of algorithms to deliver public services. Do you want to target beneficiaries of government programs and services precisely and accurately? Well, there's an algorithm for that.[1] This is also true for real-time monitoring of resources,[2] personalization of government interactions, fraud and corruption prevention,[3] anticipation (if not outright prediction) of events and behavior,[4] and more. In such instances, algorithms seem like a magic formula that can crack some of government's most persistent problems.



However, experience shows that algorithms can be divisive and destructive, be it in the hands of governments, government-affiliated partners, or forces hostile to public-sector actors. Algorithms have been used to sow distrust in public information and government machinery such as elections,[5] and they have been held responsible for perpetuating discrimination in the delivery of services and unfavorably profiling segments of the population.[6] Some have blamed algorithms for a variety of injustices, such as people being denied admission to college[7] or being denied bail by judges who rely on automated systems.[8]

With algorithms, even good intentions can result in unforeseen socially and politically disruptive outcomes. For example, when the U.K. government decided to award A-level exam grades based on an algorithm rather than actual exam results during the pandemic, almost 40% of students received lower grades than they had anticipated.[9] "F**k the algorithm" became the rallying cry of protesters who took to the streets or sought redress in court. The backlash forced the government to retract the grades. Subsequent reviews suggested that the algorithms might have been biased[10] (reinforcing prejudices in historical data and favoring smaller schools). Critics also took issue with the limited engagement and accountability tools that the government provided for students and parents.[11]

The Dutch government faced a similar reversal when in 2020, a court ruled that a digital welfare fraud detection system called *Systeem Risicoindicatie* (SyRI) was unlawful because it did not comply with the right to privacy under the European Convention of Human Rights. The law, establishing the system, had passed in 2014 without a single dissenting vote in the parliament and ostensibly contained numerous provisions to discourage 'fishing expeditions' and ensure that any harm to individuals whose data was processed by the system was proportionate to the allegations of fraud. The court however found these provisions to be inadequate and faulted the law/system on many grounds including lack of transparency, the inability to track or challenge the data, the risk of discrimination, unsatisfactory attention to purpose limitation and data minimization, and insufficient independent oversight.[12]

Complicating matters for governments, particularly those in developing countries that are eager to introduce or expand the use of algorithms in the public sector, is the fact that most of the experience and lessons learned so far reflect the reality in developed countries, where there is greater technical, human, institutional, and infrastructural capacity. What's more, advanced economies have different priorities and policy objectives than developing countries and have a different level of algorithmic maturity.

This paper presents preliminary observations drawn from a high-level review of two cases,[13] one in Izmir, Turkey, and one in Belgrade, Serbia, as well as an analysis of secondary material. The focus is on data governance-related design and implementation issues specific to developing country governments that are considering algorithmic decision-making services.

SECTION 2

Algorithms and Algorithmic Decision-Making: The Basics

What is an algorithm/algorithmic decision-making?

An algorithm is a step-by-step procedure to turn any given inputs into useful outputs. A computer algorithm follows a series of instructions to transform inputs (data) into outputs that can be used for making decisions, either by the computer system or a human. Many machine-learning algorithms learn directly from data by identifying patterns and relationships, without rules-based instructions from humans.

The algorithms discussed in this paper focus on systems that either augment or replace humans for decision-making in the public sector. One basic example is an algorithm to determine customs duty at an international border. If the value of a shipment exceeds a certain threshold, apply a duty unless exporting to a neighboring country. Determining eligibility for COVID vaccines when they are scarce is an example of a more complex algorithmic decision-making process, as it involves a greater number of variables with intricate, sometimes dynamic interrelationships.

Why is algorithmic decision-making different?

Algorithms use powers that far exceed the tools typically available to human decision-makers.[14] These include vast computing power that goes beyond human cognitive capabilities (e.g., the ability to crunch real-time data about all the vehicles on the road in a city at a point in time); constant learning without human supervision and based on patterns that are humanly impossible to discern (e.g., the ability to recognize individuals based on their gait without ever seeing their face); and dynamic nudging that creates instant incentives for compliance (e.g., a guided selection of benefits designed to promote specific economic behavior).

Is an algorithm/algorithmic decision-making the same as artificial intelligence (AI)?

The terms algorithm and AI are often used interchangeably. In policy terms, it's useful to think of algorithms—a form of automated instruction—as a subset of AI, which encompasses larger socio-political and economic issues and a variety of technical/scientific disciplines.

How do computer algorithms and humans interact?

Humans interact with algorithms as designers and creators, embedding their socio-political value systems into code; consumers and users who gain value from the code and use services/products; and as sources of data whose actions serve as new data points or inputs for the algorithm. Humans can also provide a point of control for the algorithm, either as testers or validators of the decisions made by the algorithm. Algorithmic systems that act independently, without control or supervision from humans, are considered autonomous.

SECTION 3

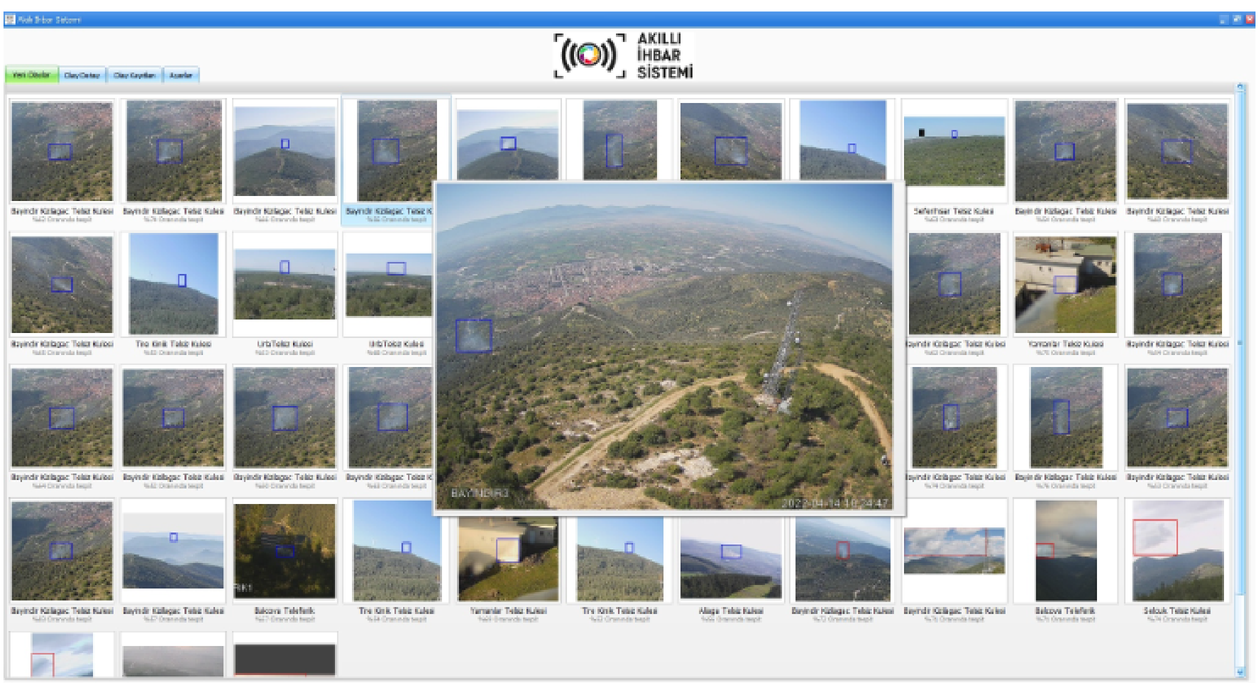
The Use of Algorithms in the Delivery of Public Services

Illustrating the Opportunity and the Challenges

Case 1: Using algorithms to combat natural hazards in Izmir, Turkey[15]

Forest fires are a constant hazard during the long, hot summers in Izmir, Turkey's third-largest city whose municipality serves an entire province of 30 districts.[16] Fires can spread quickly due to frequent high winds in the area, so it's important to detect them as soon as possible to minimize the loss of life and property.

Because it's impractical to rely on human monitors to cover the entire region, the metropolitan municipality established the Intelligent Emergency Warning System (AIS) in May 2022 to detect natural events such as fires and floods. The system relies on 45 high-resolution cameras at 12 stations covering half the territory of the Izmir municipal area. It alerts human operators when it detects any potential fire or flood incidents, and humans then validate the reports using data transmitted by the system. Then they inform the fire department and other relevant agencies and send SMS warnings to local community representatives. At the time of writing, the system had correctly identified 33 events in about four months of operation.



Data and algorithms

Cameras are the primary data source, but the system also utilizes third-party data to provide additional information to responders, such as traffic. The data is not currently available through the city's open data platform[17] because the system is still in its early stages of maturity.

Algorithms are used in multiple ways, such as:

- Processing images to accurately identify fire events
- Calculating response times
- Improving accuracy by learning from false detection

Human interaction, oversight, redressal, and stakeholder engagement

This system is not autonomous since human operators must verify and validate all incident reports. Citizens can refer to a portal or reach a call center with questions or complaints.

The system was developed and implemented by the city's in-house IT team in coordination with the local fire authority, which has operational responsibility for it. Feedback from the fire department and citizen input are used to update and maintain the system. Citizens and nongovernmental organizations can use existing feedback mechanisms to contact the implementing agencies and request follow-up.

The system was financed through the IT department's regular budget, with additional hardware costs being supported by the metropolitan municipality. Regular technical and regulatory audits are planned, as is a questionnaire-based citizen survey that will cover service awareness, service rating, and recommendations to improve service quality.

Policy/regulations/institutions

The city has deemed the existing policy/regulatory/institutional environment adequate for the system. Among the issues considered were:

- **Data protection and security:** This was handled according to the provisions of the ISO 27001 information security certification standard. There are well-defined data access mechanisms in place that are reviewed and updated periodically. An independent third-party conducts regular penetration tests. All image processing is done on site, and the data is transmitted to the emergency response center only in case of specific disaster events. The data is stored for a maximum of 30 days.

- **Privacy:** The system is considered compliant with KVKK, Turkey's personal data protection law.[18] Masking techniques are used as needed to protect the privacy of individuals and property in the live feed and for security reasons, especially when military installations are in view.

Opportunities and challenges

The system requires very high bandwidth to manage and process image data, which places a considerable burden on the city's network infrastructure. Many design decisions in the system have been influenced by the need to optimally utilize limited bandwidth. Service downtime is another challenge.

The program team considers data quality and financing (e.g., to install thermal cameras) as the primary challenges going forward. Another challenge is that most citizens are unaware of the system, despite media coverage[19] and social media campaigns. This might require greater civil society engagement.

Looking ahead

The municipality is exploring plans to extend the use of algorithms to deliver additional citizen-and business-facing services, both to improve the quality of such services and to optimize the use of resources.

Case 2: Using algorithms to augment public health infrastructure in Belgrade, Serbia[20]

As caseloads grew during the COVID pandemic, the government recognized the need to provide additional support to health care professionals to help them triage patients and continue to deliver efficient and accessible services to citizens. The Ministry of Health in Serbia has started a proof-of-concept (POC) project at multiple clinics in Belgrade to automate the reading of chest X-rays and provide initial diagnosis using an AI-based solution. The ministry selected chest X-rays for the POC because they are typically the first tool for many diagnostics involving the heart, lungs, blood vessels, airways, and even bones of the chest and spine.

The system works by having X-rays taken at different clinics sent to the Central Radiology Information System. Deep learning algorithms analyze and triage the X-rays, typically within one minute. All readings are then validated by a radiologist and compared to the original radiologist reports. At the time of writing, more than 200 images had been read by algorithm and cross-checked by radiologists. The accuracy rate of one of the algorithms being tested was 71.4%, which was lower than the vendor had claimed and lower than the solutions offered by other service providers. Results for the other algorithm were not available at the time of writing.

Data and algorithms

The data for the algorithms is sourced from multiple health care institutions. It's stored in Serbia's central healthcare information system. A third-party vendor is responsible for the management of the system and must meet contractual obligations for the security and quality of the data. The data and information about the algorithms are currently not available publicly.

Algorithms are used in multiple ways, such as:

- Processing images to provide initial diagnosis
- Triaging patients

Human interaction, oversight, redressal, and stakeholder engagement

The system is not autonomous since radiologists verify and validate all readings during the POC stage. The expectation is that the need for human verification will decline substantially after the system is put into production and there's greater confidence in the accuracy rates.

The POC has been implemented by two vendors that offer slightly different algorithms, under the supervision and operational responsibility of the Ministry of Health and the Office for IT and eGovernment. Stakeholder consultations during the design phase included health care experts, lawyers, technical experts, representatives from the Ministry of Health, and other government officials from the Office of the Prime Minister. No separate financing was provided by the government for the POC.

No technical or regulatory audits are planned for the POC. However, in line with the requirements expected to be established by the upcoming guidelines of trustworthy use of AI, the government plans to conduct a questionnaire-based assessment of trustworthy AI that will also include the stakeholders of the current POC.

Policy/regulations/institutions

Health care has been identified as a priority sector in both the Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025, as well as the accompanying action plan for 2020-2022. The strategy envisions a new Agency for the National AI Program, which is expected to be established shortly. Serbia is also developing a Law on Ethical Use of Trustworthy AI that will provide the legal framework for the current initiative as it's scaled beyond the POC stage. Meanwhile, the government is developing separate guidelines for trustworthy AI.

The government has deemed the existing policy/regulatory/institutional environment adequate for the POC. Among the issues considered were:

- Data protection and security – The contractual responsibility of the vendor managing the central healthcare information system
- Privacy – All data is anonymized

Opportunities and challenges

The system is not applicable in all patient contexts. For example, the system can't analyze data from patients who are unable to lie on their right side (PA) and left side (AP) during radiography. The project team has also identified compliance with the law and ethics as potential issues.

The project team considers data quality, financing, and an underdeveloped regulatory and institutional apparatus to be the main challenges going forward. It's also important to ensure that all AI applications within the country conform to EU standards.

Looking ahead

While the POC proceeds, the state-owned Institute for Artificial Intelligence Research and Development of Serbia is developing a similar solution that may replace the algorithms currently deployed by the vendors. Meanwhile, the government plans to extend the use of algorithms to other services, such as CT and MRI scanning, to improve the quality of services and optimize the use of resources. The preliminary results of the POC suggest that similar solutions may be effective for additional diseases, such as rectal and prostate cancer.

Emerging common questions

The cases above are just two examples of the growing use of algorithms by governments in developing countries, many of which will have profound implications for the socioeconomic well-being of people. A recent paper[21] examined emerging examples across Latin America, including the use of algorithms in policing software to predict crimes in Uruguay and to evaluate at-risk youth in Argentina. For the Argentinian initiative, the government collected data from 200,000 people living in vulnerable areas through NGOs and then developed a machine-learning model to generate predictions about school dropouts and teenage pregnancy.

Examples abound from other regions as well. In Kenya, the government recently announced a plan[22] to use algorithms to allocate affordable houses. In South Africa, different government agencies have used a locally developed platform for a range of surveillance-related activities, including policing and poaching prevention in national parks. More controversially, algorithms have been used in different countries to profile segments of the population or monitor refugees and other marginalized populations.

While examples proliferate, there has not yet been a comprehensive assessment of the quality and impact of most of the initiatives described above. However, several studies, focused mostly on developed countries are underway, including this one.[23] Much is still to be learned about these examples, but common questions are emerging, including:

- **People/social mandate:** How well-informed are people about the role of algorithms in delivering services, such as diagnoses in Belgrade or hazard detection in Izmir? How did they provide consent? What is their level of satisfaction? What redress tools are available to them? Did the selection of services demonstrate bias against certain populations? Were they involved in the design of the system?

- **Data:** What training data was used to develop the algorithm? Does the algorithm work as effectively on the local population as it does on the training data? Is the use of this data purpose limited? What are the enforcement mechanisms? How effective are these enforcement mechanisms? Should this data be available under controlled circumstances to third parties such as developers, entrepreneurs, and civil society organizations?

- **Regulations:** Is there or should there be regulatory requirements for approval before commercial solutions based on pilot algorithms are scaled up? Do existing regulations adequately address relevant privacy concerns? Are existing regulations culturally appropriate? What are the ethical questions raised by the solution/approach?

- **Infrastructure:** Does the city or country possess adequate technical infrastructure to scale the solution? Should this infrastructure, including algorithms, be open to third parties, including citizens?

- **Impact/effectiveness:** Has the pilot achieved its goals? Are these goals equitable? What new risks has the pilot introduced to either the state or individual citizens?

SECTION 4

Data Governance Issues for Developing Countries

Developing countries face many of the same data governance issues[24] that advanced economies do, and these issues are typically addressed as part of an overall digital transformation plan. However, there are a number of unique data governance issues that have greater relevance in developing countries and have a direct impact on the selection, design, and implementation of specific algorithmic decision-making initiatives by governments.

- **Issue #1:** Institutions in developing countries have an extreme legitimacy, accountability, and transparency problem.
- **Issue #2:** Poor local data means that the people of developing countries are inadequately represented in training data.
- **Issue #3:** People in developing countries have less experience interacting with machines and algorithms, and there's limited data in local languages to close the cultural gap.
- **Issue #4:** Developing countries have had limited involvement in developing standards for fairness, transparency, and accountability in algorithmic decision-making.
- **Issue #5:** Developing countries are dependent on international data infrastructure to develop and manage their algorithms.
- **Issue #6:** Developing countries deploying algorithmic decision-making are dependent on big tech companies but have little leverage over them.

Issue #1: Institutions in developing countries have an extreme legitimacy, accountability, and transparency problem.

Governments, even in advanced economies, recognize and grapple with the challenges of legitimacy, accountability, and transparency of algorithms. Part of the challenge is technical. Algorithms, given their utilization of vast computing power and their self-directed learning abilities, are inherently difficult to audit, making it hard to trace their biases.[25] Other challenges are organizational and social. A recent paper[26] provides a useful summary of the universal trust questions that inevitably accompany algorithmic decision-making (e.g., disenfranchisement, disconnection, low traceability and explainability, bias, poor quality, and reinforcement of power inequalities) and proposes a helpful trust framework that outlines legal mandates and guidelines that governments should consider.

The trust and legitimacy issues in developing countries cut deeper than in advanced economies, which tend to have a longer tradition of accountability in government and a civil society with greater power to interrogate government decisions. For example, a recent study[27] of Kenya, India, Nigeria, South Africa, and the Philippines found that existing institutions in these countries, despite formal powers, routinely fail to protect against discrimination. Another recent report[28] found that in South Africa, algorithmic scoring technologies have “deep historical roots in racist social control” and “contemporary South Africa… presents an especially stark illustration of …the ‘New Jim Code.’” While the specific findings of the studies may be contestable, they do echo many prevalent views. Many developing countries have similar colonial legacies, and it’s probable that their algorithmic decision-making apparatus is, knowingly or not, informed by discriminatory power systems (e.g., male, gendered, white, heteronormative, powerful, and Western).[29]

Policymakers in developing countries should ensure that their algorithmic decision-making is done by governmental and civil institutions that are well-rooted in a culture of transparency and statistical analysis of the disparities faced by protected groups; include vigilant nongovernmental actors attentive to algorithmic decision-making; and support a reasonably robust and proactive executive branch or an independent office to police discrimination.[30] The proposal to establish a new Agency for the National AI Program in Serbia is a welcome step, as are plans in both Serbia and Turkey to conduct regulatory and technical audits of their implementations, but the challenges described above cannot be underestimated.

Local participation is another antidote to trust and legitimacy challenges, but it’s unfortunately a known blind spot in the implementation of many algorithms. For example, this study[31] found little evidence of affected populations playing a significant role in the design or management of algorithms in the humanitarian sector.

Issue #2: Poor local data means that the people of developing countries are inadequately represented in training data.

Algorithms are not one size fits all, and wrong assumptions about algorithms can have highly consequential outcomes. This is especially true when algorithmic solutions based on evidence drawn from population studies in advanced economies are applied in developing countries.

Algorithms that purport to read medical images, as in the POC in Serbia, are one case in point. Studies have found that the data behind these types of algorithms is typically drawn from a very narrow pool, often just a single hospital. A recent Korean study[32] found that only 6% of 516 reported studies tested their algorithm at more than one hospital. Very few of these studies were conducted in developing countries or considered the characteristics of their population.

The accuracy rate of the algorithms can drop significantly in different medical settings, depending on the characteristics of patients as well as extraneous factors like the brand of equipment used.[33] Implementers in developing countries must be extremely cautious when adopting off-the-shelf algorithms that may not have taken their local population characteristics into account. The state-owned Institute for Artificial Intelligence Research and Development of Serbia is developing its own algorithm to replace those currently deployed by vendors.

Issue #3: People in developing countries have less experience interacting with machines and algorithms, and there’s limited data in local languages to close the cultural gap.

The science and art of human-machine interaction is evolving, and humans are still learning to work with machines. In the Izmir case study, machines and human operators form a team, sharing workflows to achieve a common goal. Nontech people, particularly in advanced economies, are gradually becoming more accustomed to interacting with machines in their daily lives (e.g., robot vacuums, semi-autonomous vehicles, robots on factory floors, and digital assistants like Siri and Alexa) mostly without understanding any aspect of the black box algorithms behind the machines. This inscrutability cuts both ways, and machines, whose problem-solving techniques are fundamentally different from humans, can struggle to understand the socio-cultural and ritual aspects of working with humans.[34]

In developing countries, algorithms that don't account for local cultural nuances or are deployed in populations unused to algorithmic decision-making can be particularly harmful. Algorithms that rely on machines that can't converse in local languages can make the divide even greater. At the time of writing, Google Home didn't support Zulu,[35] which is widely spoken in South Africa, one of the more developed markets in Africa.

One way to train machines and algorithms to work better with humans is to expose them to a sufficiently large corpus of commonsense knowledge informed by cultural practices. This knowledge can be either "declarative" (i.e., stop at a "Stop" sign with specific visual features) or "procedural/conventional" (i.e., do not go to the back of the store to pick up your package, wait for the store staff to bring it to you). Typical sources of such knowledge include written, video, and audio material (e.g., books, articles, movies, and cartoons[36]), ideally online in digital format.

The quantity and quality of available explicit knowledge about developing countries is relatively low, and even lower in local languages. For example, according to one estimate,[37] 60% of the 10 million most popular websites on the internet are in English. Hindi, spoken by more than 600 million people worldwide, is the top South Asian language but accounts for only 0.1% of online content. Other languages like Bengali and Urdu, which are spoken by hundreds of millions of people, don't even appear on the list. Content in the African language of Igbo, spoken by at least 30 million people, makes up less than 0.1% of all online material. Initiatives such as Masakhane,[38] Zindi,[39] and No Language Left Behind[40] are steps toward addressing the issue, but the chasm remains wide.

Natural language processing implementations, like the one in Serbia, must consider that algorithms may not account for structural differences between languages, and machines may inadvertently become trained to perpetuate stereotypes. For example, Turkish does not have a gender pronoun. So when some machines translate the word "cook" into English, they identify the cook as a woman, while assigning the male gender to professions such as doctor and engineer.[41]

Issue #4: Developing countries have had limited involvement in developing standards for fairness, transparency, and accountability in algorithmic decision-making.

As governments have turned toward algorithmic decision-making, issues such as fairness, transparency, and accountability have increasingly come to the fore. Western countries have been early to respond, and several governments and independent organizations have developed guidelines, charters, laws, and regulations designed to ensure that algorithmic decision-making is equitable and inclusive. Notable examples include the Algorithm Charter of New Zealand;[42] the Ethics, Transparency and

Accountability Framework for Automated Decision-Making in the United Kingdom;[43] and the Digital Republic Law in France.[44] Many developing countries have followed suit. For example, Uruguay,[45] India,[46] and Tunisia[47] have developed strategic approaches to AI that contain many provisions for algorithmic fairness and transparency in line with the Western model.

Civil society actors have raised concerns about whether Western ideas of fairness should be considered universal and if they apply unquestionably in developing countries. Advanced economies have legal traditions based on enlightenment values, ideas of structural injustices largely centered on race and gender, and AI tools based on datasets like ImageNet (one of the most widely used training datasets in the world), reflect many Western biases.[48] Some scholars have questioned the primacy of Western ethical traditions in most AI systems and wondered whether incorporation of ethical beliefs based on alternative systems inspired by Buddhism, Shinto, or Ubuntu, for example, might change some assumptions about ethical AI.[49]

A study on algorithmic fairness in India[50] identified three factors that policymakers should focus on there, which might also be relevant in many other developing countries: 1) Data and model distortions that privilege wealthy, mostly middle-class men and minimally represent local structures like caste and sub-caste, indigenous Adivasis, and social justice practices like job reservations; 2) Algorithm designers who take advantage of poor redressal avenues available to marginalized people, using these populations as Petri dishes for intrusive practices that might not pass muster in other geographies; and 3) Unquestioning belief in positive and fair outcomes through AI without creating an ecosystem of actors to help achieve them.

The challenges for developing countries are compounded by an international AI regulatory/policy ecosystem that is largely dominated by developed countries. China and India have a growing voice in international institutions and bodies considering AI-related standards and guidelines, but most developing countries are underrepresented in these institutions, as demonstrated by the chart below.[51]

	State-led AI governance	Non-state-led AI governance
EMBEDDED IN EXISTING ARCHITECTURE	• G7 • G20 • CCW Group of Governmental Experts on emerging technologies in the area of LAWS (GGE) • Council of Europe (CoE)	• United Nations • European Commission • Organization for Economic Co-operation and Development (OECD) • IEEE • ISO/IEC
ESTABLISHING NEW INSTRUMENTS	• Global Partnership on AI (GPAI) • AI Partnership for Defense	• Partnership on AI (PAI)

Figure 1: Types of Governance and Institutions

Issue #5: Developing countries are dependent on international data infrastructure to develop and manage their algorithms.

The deployment of algorithms at scale is resource intensive, requiring large amounts of data and a highly sophisticated and expensive computing infrastructure. According to one estimate,[52] it can cost upward of $150,000 to train a contemporary neural network for an English to German translation engine, and that network would release emissions equivalent to a trans-America flight. The costs alone make the introduction of algorithmic decision-making a daunting proposition in most developing countries.

In Izmir, the implementation team had to make many design decisions to account for the local network not being able to support the bandwidth required by the city's algorithm. The city also lacked the financing required to install additional thermal cameras that would increase the effectiveness of the algorithms.

Developing countries typically don't have the complex infrastructure of data storage and modern computing hardware required to test and run algorithms, so they are dependent on infrastructure provided by large firms based in foreign countries. Compounding the problem is the fact that the global data storage infrastructure is unevenly distributed. One study estimated that the United States accounts for almost 40% of all global data storage sites,[53] with another five countries accounting for an additional 30%. California alone has more data centers than all of sub-Saharan Africa.[54] Amazon, Google, and Microsoft manage more than 50% of the world's data centers, while Chinese firms operate the world's largest ones. This leaves developing countries in a tenuous position, especially as concerns about data localization and sovereignty[55] mount and regulations around the transfer of personal data across national boundaries become restrictive.

While the digital strategies of many developing countries such as Nigeria[56] and Vietnam[57] include provisions for data centers and cloud services, they are often limited to government data. Therefore, most developing countries are at the mercy of international operators.[58]

Issue #6: Developing countries deploying algorithmic decision-making are dependent on big tech companies but have little leverage over them.

As described above, large international firms still control access to the computing infrastructure and data required to develop, manage, and implement algorithmic decision-making in most countries. The impact of this dependence is worse for developing countries, whose generally low-per capita income and, in many cases, small size mean they have little leverage over these large firms.[59] Contrast this with the situation when Europe implemented its General Data Protection Regulation (GDPR). When that happened, companies throughout the entire market scrambled, often at great cost, to update their digital products, services, and conditions to meet the requirements of the GDPR because they did not want to lose access to some of the world's largest economies.

Very little work has been done to rigorously examine the effects of the power imbalance between developing countries and large international digital firms and platforms. As noted above, some countries have responded with a pastiche of disjointed approaches, such as bans, social media taxes, and data localization requirements, but there's limited agreement on more positive responses, such as regional data pools and shared computing infrastructure.

SECTION 5

The Way Forward for Developing Countries

This issue brief outlines some of the challenges developing countries face in designing and implementing algorithmic decision-making tools at scale. The case studies presented are about projects at an early stage of implementation. It's likely that as other cases are considered, designed, and implemented in more contexts and settings, new lessons will emerge. Meanwhile, a few action steps to consider include:

- **Create regional or other alliances to tackle relevant data governance challenges.** The EU is an example of this, but countries may also consider alliances that are not based on geography.
- **Focus on cases that don't depend on personal data to deliver relevant services to citizens and businesses.** The fire detection system in Turkey is an example of this, but there are many other infrastructure management and business service possibilities.
- **Keep the focus on people, since all algorithms affect people directly or indirectly.** Governments must develop and implement engagement strategies that are designed to be inclusive and continuous, as well as to recognize the primacy of people as designers and supervisors of algorithms and as consumers of algorithmic services.

An upcoming paper from DIAL will focus on specific operational tools and resources that developing countries may consider.

Endnotes

1 "Machine learning and phone data can improve targeting of humanitarian aid," https://www.nature.com/articles/s41586-022-04484-9.

2 "Real-Time asset tracking; A starting point for Digital Twin implementation in Manufacturing," shorturl.at/CEKOX.

3 "Using AI and machine learning to reduce government fraud," https://www.brookings.edu/research/using-ai-and-machine-learning-to-reduce-government-fraud/.

4 "The new science of sentencing," https://www.themarshallproject.org/2015/08/04/the-new-science-of-sentencing.

5 "'The algorithm has primacy over media … over each of us, and it controls what we do,'" https://hls.harvard.edu/today/the-algorithm-has-primacy-over-media-over-each-of-us-and-it-controls-what-we-do/.

6 "Racism in, racism out: A primer on algorithmic racism," https://www.citizen.org/article/algorithmic-racism/.

7 "The death and life of an admissions algorithm," https://www.insidehighered.com/admissions/article/2020/12/14/u-texas-will-stop-using-controversial-algorithm-evaluate-phd.

8 "Policy brief: Pretrial algorithms (risk assessments), https://bailproject.org/wp-content/uploads/2022/07/RAT_policy_brief_v3.pdf.

9 "A-level results: Almost 40% of teacher assessments in England downgraded," https://www.theguardian.com/education/2020/aug/13/almost-40-of-english-students-have-a-level-results-downgraded.

10 "Awarding GCSE, AS & A levels in summer 2020: Interim report," https://www.gov.uk/government/publications/awarding-gcse-as-a-levels-in-summer-2020-interim-report.

11 "'F**k the algorithm?: What the world can learn from the UK's A-level grading fiasco," https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/.

12 "Social Welfare, Risk Profiling and Fundamental Rights: The Case of SyRI in the Netherlands," https://www.jipitec.eu/issues/jipitec-12-4-2021/5407#:~:text=The%20Court%20held%20that%20SyRI,basis%20of%20Article%208%20ECHR.

13 Please note that the case study descriptions are based on material provided by the teams working on the projects. They have not been analyzed critically. A future report from this author will include the perspective of additional stakeholders and deeper analysis of the accompanying systems.

14 "AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings," https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7164913/.

15 Based on inputs provided by the Government of the Izmir Metropolitan Municipality IT Department and the Fire Emergency Department.

16 "Izmir Metropolitan Municipal Authority Strategic Plan 2015-2019," https://www.izmir.bel.tr/CKYuklenen/EskiSite/file/MALI_HIZMETLER/StrategicPlan2015-2019.pdf.

17 "Izmir Metropolitan Authority open data platform," https://acikveri.bizizmir.com.

18 "KVKK (Personal Data Protection Law of Turkey)," https://www.kvkk.gov.tr/Icerik/6649/Personal-Data-Protection-Law.

19 "Protection shield for forests in Izmir," https://www.cumhuriyet.com.tr/turkiye/izmirde-ormanlara-koruma-kalkani-1950805.

20 Based on inputs provided by the Government of the Republic of Serbia.

21 "Algorithms and artificial intelligence in Latin America," http://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Screen_AW.pdf.

22 "Affordable housing program purchase allocation criteria," https://bomayangu.go.ke/downloads/20200608_AHP_Allocation_Criteria.pdf.

23 "Algorithmic accountability for the public sector," https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/.

24 Apart from the questions noted in the main body of the report, it is important to reiterate that developing countries seeking to introduce or expand the use of algorithms in government face many of the same data governance challenges that governments in advanced economies do, and that inevitably accompany digital transformation everywhere. The following is an illustrative list, but is not the focus of this brief: the data and computing infrastructure required to run sophisticated algorithms tends to be inadequate; the quality of data—its completeness, biases, frequency, coverage, and access—are common problems; the legitimacy, accountability, and transparency questions about algorithms are difficult to resolve; the regulatory and institutional environment to tackle the issues raised by the proliferation of algorithms have generally not kept pace with some complaining that regulations stifle innovation, entrepreneurship, and competition and others grousing that institutions don't sufficiently protect the interests of the weak, the marginalized, and the vulnerable; skills, capacity, and participation gaps are recurring themes within the government, civil society, private sector, and academia; and financing is never adequate or timely.

25 "The hidden dangers in algorithmic decision making," https://towardsdatascience.com/the-hidden-dangers-in-algorithmic-decision-making-27722d716a49.

26 "A trust framework for government use of artificial intelligence and automated decision making," https://arxiv.org/pdf/2208.10087.pdf.

27 "What does automated decision-making portend for the fight against discrimination in developing countries?" https://digi-con.org/what-does-automated-decision-making-portend-for-the-fight-against-discrimination-in-developing-countries/.

28 "Apartheid by algorithm," https://logicmag.io/home/apartheid-by-algorithm/.

29 "Machine ethics and African identities: Perspectives of artificial intelligence in Africa," https://www.researchgate.net/publication/361644515_Machine_ethics_and_African_identities_Perspectives_of_artificial_intelligence_in_Africa.

30 "Algorithmic decision-making and discrimination in developing countries," https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=1135&context=jolti.

31 "Predictive analytics in humanitarian action: A preliminary mapping and analysis," https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/15455/EIR33_Humanitarian_Predictive_Analytics.pdf?sequence=1&isAllowed=y.

32 "Design Characteristics of Studies Reporting the Performance of Artificial Intelligence Algorithms for Diagnostic Analysis of Medical Images: Results from Recently Published Papers," https://pc.kjronline.org/DOIx.php?id=10.3348/kjr.2019.0025.

33 "Artificial intelligence could revolutionize medical care. But don't trust it to read your x-ray just yet," https://www.science.org/content/article/artificial-intelligence-could-revolutionize-medical-care-don-t-trust-it-read-your-x-ray.

34 "Human-centered artificial intelligence and machine learning," https://arxiv.org/abs/1901.11184.

35 "Conversational AI: Africans disproportionately disadvantaged," https://www.context.news/ai/opinion/conversational-ai-africans-disproportionally-disadvantaged.

36 "CIDEr: Consensus-based image description evaluation," https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Vedantam_CIDEr_Consensus-Based_Image_2015_CVPR_paper.pdf.

37 "Usage statistics of content languages for websites," https://w3techs.com/technologies/overview/content_language.

38 "Masakhane: A grassroots NLP community for Africa, by Africans," https://www.masakhane.io.

39 Zindi website, https://zindi.africa.

40 "No Language Left Behind," https://ai.facebook.com/research/no-language-left-behind/.

41 "He is a doctor, she a nurse: How language carries gender bias into algorithms, perpetuates status quo," https://www.outlookindia.com/culture-society/he-is-a-doctor-she-a-nurse-how-language-carries-gender-bias-into-algorithms-perpetuates-status-quo-news-195387.

42 "Algorithm charter for Aotearoa New Zealand," https://data.govt.nz/toolkit/data-ethics/government-algorithm-transparency-and-accountability/algorithm-charter/.

43 "The Ethics, Transparency and Accountability Framework for Automated Decision-Making," https://www.gov.uk/government/publications/ethics-transparency-and-accountability-framework-for-automated-decision-making.

44 "Digital Republic Law," https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033205514/.

45 "Artificial Intelligence Strategy for the Digital Government," https://wp.oecd.ai/app/uploads/2021/12/Uruguay_Artificial_Intelligence_Strategy_for_Digital_Government_2019.pdf.

46 "National Strategy for Artificial Intelligence," https://indiaai.gov.in/research-reports/national-strategy-for-artificial-intelligence.

47 "National AI strategy: Unlocking Tunisia's capabilities potential," http://www.anpr.tn/national-ai-strategy-unlocking-tunisias-capabilities-potential/.

48 "Excavating AI: The politics of images in machine learning training sets" https://excavating.ai.

49 "Ethically aligned design," https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf.

50 "Re-imagining algorithmic fairness in India and beyond," https://arxiv.org/pdf/2101.09995.pdf.

51 "Mapping global AI governance: A nascent regime in a fragmented landscape," https://link.springer.com/article/10.1007/s43681-021-00083-y.

52 "Energy and policy considerations for deep learning in NLP," https://arxiv.org/pdf/1906.02243.pdf.

53 "Microsoft, Amazon and Google Account for Over Half of Today's 600 Hyperscale Data Centers," https://www.srgresearch.com/articles/microsoft-amazon-and-google-account-for-over-half-of-todays-600-hyperscale-data-centers.

54 “Improving data infrastructure helps ensure equitable access for poor people in poor countries,” https://blogs.worldbank.org/opendata/improving-data-infrastructure-helps-ensure-equitable-access-poor-people-poor-countries.

55 “Sovereignty and Data Localization,” https://www.belfercenter.org/publication/sovereignty-and-data-localization.

56 “National Digital Economy Policy and Strategy (2020-2030): For a Digital Nigeria,” https://www.ncc.gov.ng/docman-main/industry-statistics/policies-reports/883-national-digital-economy-policy-and-strategy/file.

57 “National strategy for development of digital economy and digital society to 2025, orientation to 2030,” http://www.asemconnectvietnam.gov.vn/default.aspx?ZID1=3&ID1=2&ID8=118296.

58 “Developing countries are being left behind in the AI race—and that’s a problem for all of us,” https://theconversation.com/developing-countries-are-being-left-behind-in-the-ai-race-and-thats-a-problem-for-all-of-us-180218.

59 “Governing big tech’s pursuit of the ‘next billion users,’” https://www.cgdev.org/sites/default/files/governing-big-techs-pursuit-next-billion-users.pdf.