DCAF Geneva Centre for Security Sector Governance

Artificial Intelligence and SSG/R

ISSAT Advisory Note

June 2023

Acronyms

AI

SSG/R

Executive Summary

The past decade, and particularly the COVID-19 period, have seen a sharp increase in the use of digital capabilities in the area of security and justice, through e-governance practices, surveillance technologies, Big Data practices, and increased cyber security measures.

The growing use of Artificial Intelligence (AI) is part of this trend and is very relevant for Security Sector Governance and Reform (SSG/R). AI systems are already helping to improve public service delivery effectiveness. They have also allowed the generation of data-based forecasts and the simulation of complex scenarios to support decision-making.

However, these technologies can also be misused, and therefore have the potential to reduce public trust in the security and justice sector due to access and awareness shortcomings and offer privacy and bias problems.

With these challenges in mind, the objectives of SSG/R remain entirely relevant to a world where AI is shaping security and justice sectors. Good security sector governance shapes security sectors that are people-centric, respectful of human rights, effective and accountable. Building public awareness and engagement, developing capabilities and ensuring inclusiveness and respect for human rights and supporting the development of checks and balances systems are all necessities for this growing field.

Today, many states are already using technology-driven solutions in their security and justice sectors , including in the areas of access to justice, cyber-security, defence reform strategy design, legal frameworks' development, e-governance and open data initiatives. These cases provide useful insights into where the field may go.

In this note, ISSAT focuses on AI and the opportunities it creates for reform, as well as the challenges. It provides a structure for clarifying the links between AI and SSG/R as well as, present some of the emerging lessons and good practice from field-level, donor programming.

Definition of AI

AI is computer software that imitates the way that a human thinks in order to conduct complex tasks such as reasoning and analysis. There are various approaches to achieve this intelligence, including Machine Learning (ML)[1] and Natural Language Processing (NLP).[2]

AI can play a significant role in improving security and justice services delivery, as well as governance. It can provide the data necessary for evidence-based decision-making, resource allocation, as well as improving access to justice.

[1] Today, ML is the most impactful subset of AI. Broadly, it is an approach that allows computers to learn without being explicitly programmed. This is done by 'training' the computer model on large amounts of data, which it uses to train itself to find patterns.

[2] NLP is software that allows computers to understand 'natural' languages (e.g. Spanish, Mandarin) and interpret human communication.

The Challenges of AI to Justice and Security

The concept of AI has long provoked fear and fascination, and stories of AI in the service of security and defence are in fact millennia old.[3] Whether or not our vision of AI is more about destruction or protection, AI offers many critical challenges for justice and security. Among the most discussed today are:

- **Perpetuation of bias and discrimination:** AI systems are predominantly based on data and are as good as the data they use. If data reproduces systemic discrimination, then so will the technology, even if this is unintentional.
- **Power play:** AI can be used as an instrument of political oppression, provocation, persecution and manipulation, if used without adequate checks and balances, including through surveillance, deep fakes and other disinformation, and political agenda setting.
- **Regulation playing catch up:** In many contexts, the use of technology in the security and justice sector is happening before legal frameworks have been adequately developed. This can undermine civic, consumer and human rights. Generating new norms, policies and laws is difficult because the technology is complex and changing quickly.

AI and SSG/R intersect particularly in key programmatic areas: **Cybersecurity**, **Rule of Law and Governance** and **inclusive reform through technology** (or the use of technology to promote oversight, inclusivity and accountability). Good security sector governance will contribute substantially to the mitigation of the risks of AI to human rights, security and safety. It does so through governance and Rule of Law reforms, creating oversight and accountability mechanisms, providing the security of data platforms, as well as the usage of technology to make reforms more inclusive.

AI systems require proper legislative and regulatory frameworks in place to secure private data, to scrutinise government and private sector activities and to prohibit poor implementation which could impair fundamental rights.

Five Ways AI can support SSGR

AI can improve the effectiveness of public service delivery by raising productivity while also reducing operating costs. Not only do digital solutions provide more accessible and more efficient processes, but they also allow better information management through centralising data acquisition and management.

The exponential growth in the usage of AI is transforming the security and justice sector. AI supports preparedness and resilience in times of crises, such as pandemics, and countries with more robust technological solutions prior to COVID-19 were generally more resilient during the pandemic.

AI can improve transparency in the security and justice sector. For instance, following up on corruption cases can be longwinded, data heavy and difficult to manage. AI tools can help make sense of the information.

Public-private partnerships have a lot to offer. Cooperation between public and private actors underpins will underpin the expansion of AI in the security and justice sectors. Technologies supplied by the private sector will be applied to the data typically handled by the public sector. To maintain the legitimacy of this process, it is essential to involve Rule of Law institutions, civil society, academia and citizens in decision-making around the type of data to be collected and data protection mechanisms.

AI can enable data-driven reforms by increasing in capturing and storing data, thus generating more data-driven reporting on needs, as well as accessibility and use patterns. This system better enables data-based governance on the medium and long term.

What could the International Community Do Better?

1. **Support the development of national Rule of Law Frameworks**, which can be adapted to new AI realities. AI can provide far-reaching powers to the security sector and countries with weak Rule of Law

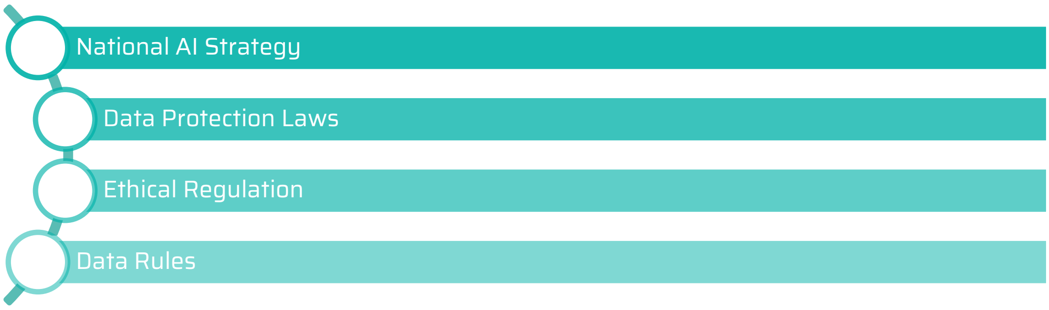
[3] Alex Shashkevich, "Stanford researcher examines earliest concepts of artificial intelligence, robots in ancient myths," 28 February 2019, https://news.stanford.edu/2019/02/28/ancient-myths-reveal-early-fantasies-artificial-life/

frameworks present the highest risks of excesses and abuses by the security sector due to gaps in the mechanisms of checks and balances.

2. **Provide international best practice** on how to integrate technology in the provision and the oversight of security and justice services. In many cases, AI applications and systems are provided to governments through the private sector or through donor assistance. There might be very few capabilities available with national authorities to handle the ethical, legal and communications aspects around data management.
3. **Develop Operational Guidance** around modalities of collecting and managing personal data for government-set needs. Transparency over government activities can be greatly enhanced by digital technologies. However, communicating to community, that personal data can become accessible and usable by the security sector is crucial.
4. **Strengthen the Capacity of State Actors** for the procurement of AI applications, usage, maintenance and sustainability. The use of new technologies in public institutions inevitably creates a skill gap. The current workforce will need to become familiar with new work practices and methodologies to guarantee that technological use is effectively optimised and sustainable.
5. **Facilitate the development of public-private partnerships** geared towards the development of technology and AI systems. International partners have a key role to play in developing public-private partnerships, building on the resources available in the private sector while ensuring that technology is developed in accordance with public standards.
6. **Encourage and guide national authorities towards Whole-of-Government and Whole-of-Sector approaches** to increase public digital coordination and collaboration. AI solutions should be applied through cross-government coherence so as not to negatively impact institutional cross-collaboration. This is also a useful practice to encourage sectoral transformation.
7. **Provide technical advice on how to reach the people who are often 'transparent' to the system.** Those who have been marginalized by the mainstream systems, are often less-educated, suffer from identity-based discrimination, socioeconomic inequality, or lack of awareness or willingness to engage with the system. Technology could further alienate these groups. International partners could support the financing of and roll-out of studies such as: community-based needs assessments, stakeholder analysis, gap analysis, gender analysis and conflict analysis. All of which should give a better reflection of who needs security and justice services and how best to close their accessibility gap with the system.

International Support to National Strategy and Rule of Law

National strategies and legal frameworks should shape and promote the development of AI systems enhancing overall governance, accountability mechanisms and coordinate public-private partnerships in this area.



National AI Strategy

Data Protection Laws

Ethical Regulation

Data Rules

National AI Strategy

A national AI strategy is a policy framework, channelling a country's national resources towards the development of its AI capability, including clear commitments for updating and rolling-out corresponding checks and balances system. It should have a clear and realistic purpose around how AI systems' benefits can be harnessed, whilst promoting a culture of accountability. It identifies what needs to change and what is the purpose of change. This vision is necessary to build a clear and effective policy roadmap, with clear targets for change. A national AI strategy should also evaluate available capital and infrastructure for the roll-out of the vision. A sectoral approach can be useful to explore the potential of AI as relevant to each sector.

A National AI strategy should include a timeline, allowing phased implementation, along with a monitoring mechanism with regular progress updates against concrete and measurable outputs.

Countries with a full (or forthcoming) National Strategy in 2021

OECD Countries	Europe	MENA	Asia	Latin America	Africa
Austria	Russia	UAE	China	Chile	Kenya
USA	Estonia	Oman	Singapore	Brazil	Mauritius
Canada	Latvia	Saudi Arabia	Taiwan	Uruguay	
UK	Serbia	Tunisia	Malaysia	Argentina	
Norway	Romania	Egypt	Indonesia	Colombia	
France	Slovenia	Israel	Vietnam	Mexico	
Australia	Czech Republic	Cyprus	India		
Sweden	Slovakia				
New-Zealand	Hungary				
Denmark	Croatia				
Netherlands	Poland				
Germany	Bulgaria				
Ireland	Ukraine				
Belgium	Lithuania				
Italy	Malta				
Portugal					
Finland					
Greece					
Spain					
Japan					
South Korea					

Key dimensions of a national AI strategy

Providing a set of standardised data-protection laws and addressing ethical concerns

Data is at the core of AI mechanisms and data-protection laws are necessary for addressing ethical concerns and regulating access to, and use of data. These laws should clearly define the relationships between data subjects and stakeholders handling data, by explaining how data is collected, stored, processed, shared and erased. Data protection laws, within the security realm, should prioritise human rights commitments made by the country, at the national and the global levels.

Establishing a strong research environment and shaping public-private integration

National strategies should enable and define a space for public-private partnerships geared towards research and development of AI systems. Although the public sector holds varied data banks as it naturally collects information to guide decision-making, it still lacks the capacity to develop and implement AI technologies on its own. However, this partnership presents risks and needs to be guided by ethical governance frameworks.

Adapting the workforce to AI technology

Any development of AI technology should anticipate an impact on the current workforce, as well as the need for future skills and infrastructure that should support the transition to AI systems. AI technology will necessitate new skills and infrastructure, and will take away certain types of jobs, via automation of previously human-run tasks. A capability building plan, as well as necessary functional reviews will should be considered in the strategy. Security and justice personnel need to be able to understand the functioning of AI systems and adequately monitor the performance of AI systems.

Engaging in international collaboration

AI technologies involve expertise and capabilities unevenly distributed across the world. In light of the pace and scope of global AI research, governments cannot work in isolation. Exchanging information and building international partnerships should support the deployment of AI systems in countries with lower ITC development capabilities. International collaboration is also key to strengthening the governance of AI according to international best practices.

Data Protection Laws and Ethical Regulation

AI Ethical regulation defines a system of moral principles intended to inform the development and use of AI technology. These key principles tend to revolve around the concepts of “Autonomy”, “No Harm”, “Benefit” and “Justice”. These concepts are central to establishing adequate auditing frameworks which will oversee the accountability of AI technology.

Autonomy
No Harm
Benefit
Justice

Autonomy

Implementing AI applications means relinquishing human decision-making power to technological solutions. The key question that arises from this transition is: 'how much?'. The balance between human-retained decision-making power and delegation to machines is at the core of the principle of “autonomy”. The goal is to avoid that increased artificial autonomy impedes the flourishing of human autonomy. To achieve this end, AI systems need to be developed and deployed with the notion that human autonomy should be promoted, and that machine autonomy should be restricted and reversible, should there be a need to re-establish human decision-making power.

Principles associated with the concept of “Autonomy”

Power to Decide	Human control	Human oversight	Transparency
Explainability	Explicability	Liberty	Openness
Fundamental rights	Human values	Personal privacy	Privacy protection

Benefit

The concept of “benefit” in AI ethics articulates the need to “prioritise human well-being as an outcome in all designs.”[4] AI technology should guarantee the basic preconditions for life on the planet, continued prosperity for mankind and environmental preservation for future generations.

Principles associated with the concept of “Benefit”

Promoting well-being	Benefit society	Generate net benefits
Sustainability	Impact	Efficacy
Explicability	User-centred design	People-first approach

No Harm

The concept of No Harm stresses the need for safeguards protecting individuals’ fundamental rights against the overuse and the misuse of AI technologies. The prevention of infringements on personal privacy is a core dimension of this concept. It includes other provisions around other threats, such as a potential AI arms race or the recursive self-improvement of AI systems. The concept of “No Harm” covers both the people developing AI and the technology itself.

Principles associated with the concept of No Harm

Risk Control	Safety	Security
Capability Caution	Data Protection	Privacy (to avoid harm)
Explicability	Transparency (to avoid harm)	Reproducibility
Accuracy	Reliability	Responsible deployment

[4] The IEEE Initiative on Ethics of Autonomous and Intelligent Systems (2017). Ethically Aligned Design, v2. https://ethicsinaction.ieee.org

Justice

The concept of “Justice” addresses the consequences of disparities in decision-making capacity within any country. It ensures that AI contributes to global justice, equal access to benefits, shared prosperity, and the elimination of all types of discrimination. It aims to mitigate the risk of bias in datasets, which are heavily influenced by systemic norms.

Principles associated with the concept of Justice

Fairness	Fundamental rights	Equality	Non-discrimination
Avoiding bias	Inclusivity	Diversity	Data Neutrality
Representative Data	Shared prosperity	Social & economic impacts	Avoid disparity
Mitigating social dislocation	Preserving solidarity	Accessibility	Explicability
Transparency (For accountability)	Openness (for accountability)	Accountability	Auditability
Liability	Judicial transparency	Open governance	Regulatory & legal compliance

Levels of AI legislation

Data rules

Data rules condition the access to the raw material needed for the development of AI products and are the first level of compliance for any AI system. They govern how data can be collected, processed and transferred. Data rules often are tied to national settings. While there are examples of supranational data legislation such as the EU GDPR, often these regulations are also augmented by additional national data protection rules.

Application-Specific AI rules

Application-specific AI rules represent the second layer of compliance. These rules explicitly target specific AI applications or domains. Although they are still rare, they are likely to increase, especially in terms of technical product norms. AI rules can span from an outright ban of specific technologies to detailed technical standards regulating AI systems. They are a combination of traditional legislation and constitutional rules, backed by sanctions, in addition to newly developed technical standards by more amorphous, non-governmental actors such as standardization organizations. Despite their origins, these rules would still be de facto and de jure binding.

General AI rules

Thirdly, general AI rules cover AI functions, such as automated decision-making. For instance, the EU GDPR implemented a semi-ban on “decision-making solely based on automated processing of personal data.” The regulation considers that automated processing can be conducted only if the decision is necessary for entering/performing a contract, it is authorised by law or if the data subject provided explicit consent. It is likely that AI General Rules will grow in depth and breadth, as future legal proceedings will force the justice system to adapt to and better interact with AI technologies.

Application-specific non-AI rules

Fourthly, application-specific non-AI rules apply to specific applications without considering AI in itself. AI use does not waive regulations applied to specific sectors. As long as no specific exemption is in place, AI systems need to abide by the specific sectoral regulation. For instance, shipping regulations stipulate that a ship needs to be operated by a captain and a crew fulfilling specific human functions. As such, this rule would become a barrier to the full automation of a cargo ship transporting goods across continents.

General non-AI rules

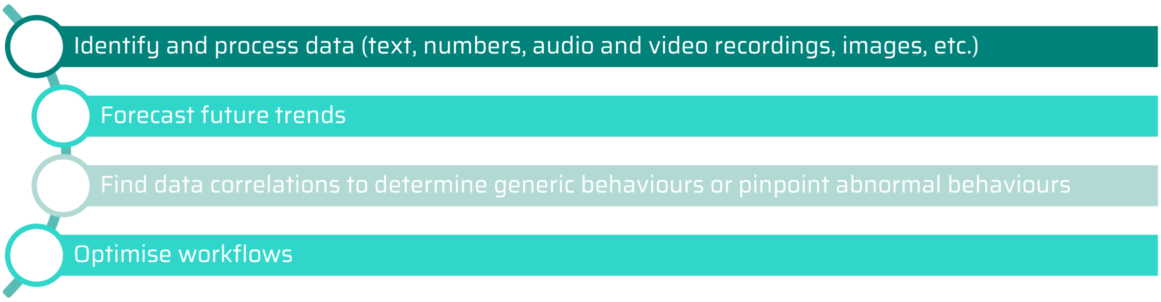
Finally, the fifth layer of compliance, general non-AI rules, provide a general legal framework for behaviour control. Typically, antidiscrimination laws can be categorised as general non-AI rules and these laws are central to the development of AI. For instance, these rules could influence algorithm design, the type of usable data, eligible decision-making parameters, and the realm of applicable decisions.

AI and the Security and Justice Sector

The application of AI in the security and justice sector is in its early days, and the utility and risks are still being understood. The Parliament of the European Union summed this up in 2021: [5]

Digital technologies in general and the proliferation of data processing and analytics enabled by artificial intelligence (AI) in particular, bring with them extraordinary promises and risks […] but also great risks for fundamental rights and democracies based on the rule of law; whereas AI should not be seen as an end in itself, but as a tool for serving people, with the ultimate aim of increasing human well-being, human capabilities and safety.

Given the great power of AI, both the risks and the opportunities should be treated with serious caution. In many fragile and conflict affected countries, security and justice institutions struggle to cover all their duties, including covering the full scope of the geographic space as well as responding to all of the human security needs of the population. This has been a source of concern for global peace and security. The lack in capacity to provide the right response to security challenges at the right time and in the right way has undermined State legitimacy and the credibility of the governance systems in place, thus challenging the social contract in place. AI could provide a renewed opportunity for reform that could be more inclusive, effective and efficient. Security and justice institutions use AI in the four below scenarios:



Identify and process data (text, numbers, audio and video recordings, images, etc.)

Forecast future trends

Find data correlations to determine generic behaviours or pinpoint abnormal behaviours

Optimise workflows

AI and the Defence Sector

In a defence context, AI offers a range of advantages:

- Quicker operational decision-making through swift access to data and its ability to identify **underlying trends** that are not easily recognized by humans, allowing better **assessment of threats and optimal responses**.
- Saving time to access and process data by using AI also provides more room for manoeuvre regarding planning and conducting operations.
- Identification of **risks factors related to working environments** and conditions and suggest protective measures to mitigate negative impacts.

[5] European Parliament, “Texts Adopted - Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial Authorities in Criminal Matters - Wednesday, 6 October 2021,” October 6, 2021, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html.

- Running **simulations** to support the training capacity of service personnel.
- In combination with robotics, helping service personnel to stay at a safe distance when dealing with contaminated environments.
- Undertaking **auxiliary and repetitive tasks** that are time-consuming for personnel. By relieving personnel from these charges, AI allows serving officers to devote more time to high value-added tasks such as strategizing and decision-making.
- Developing **forecasting models** that supports the optimisation of logistical networks, equipment management and maintenance, financial commitments, and recruitment.

AI and the Police

In a policing context, AI has a number of established uses:

- **Facial recognition** involves capturing images of the public and running a search in a pre-existing database of suspects, or indeed building a new database. This practice is either conducted live or with recorded footage. Police forces also have access to retrospective facial recognition through law enforcement databases, which contain millions of facial images.
- **Image processing** that can read lips, analyse writing styles, identify stolen cars and detect shoplifters' behaviours.
- **Financial analysis** using AI can identify criminal activities such as fraud and money laundering.
- **Predictive policing** is a practice which helps identify locations or individuals with higher risk of committing criminal activity. Crime hotspots and patterns are mapped based on historical data. This analysis feeds into crime prevention strategies, for example, police patrol organisation and other resource allocation.
- **Solvability assessment** AI can evaluate how likely it is for a case to be solved. In this instance, algorithms can recommend on resource distribution to certain cases.

AI and the Justice sector

Advanced case-law search engines use AI to facilitate information retrieval for legal cases. AI enables the creation of an **intelligent legal research system** which analyses legal precedents and provides an automated statistical analysis and summary of legal concepts used in comparable cases. AI technology can also serve as an **alternative dispute resolution mechanism**. Justice systems typically receive a substantive number of recurring low-value civil litigation cases. AI can allow complainants to go through a platform for the automatic diagnosis of the dispute, which results in solution proposals before the case is brought to court.

Predictive justice is also being researched, with the aim of developing AI systems able to predict ruling outcomes. For instance, these systems can establish whether an individual could benefit from probation by determining whether the said individual is likely to commit further crimes.

Case Studies

Access to Justice in Colombia

Developed by the University of Buenos Aires IALAB in collaboration with the Colombian Constitutional Court with the support of international donors, the PretorIA programme was launched by Colombia in mid-2020, with the aim of accelerating and improving justice provision. It allows citizens to receive immediate protection against the violation of fundamental rights through the project *Acción de Tutela* (Constitutional Action for the protection of fundamental rights). The Acción de Tutela sets legal precedents for the provision of fundamental rights based on the priority cases – key tutelas – it receives daily. With almost 3000 tutelas received each day – and factoring in the average time it takes for a person to read, analyse, and systematise its content (36 minutes) - working through all the tutelas is not feasible. By digitalising the necessary analysis and providing sentencing predictions based on predetermined criteria, PretorIA allows for more efficient and rapid decision-making.

The AI application generates intuitive reports and statistics which accelerates the procedure, thereby enabling a more rapid deployment of this legal protection to Colombian citizens. The integration of AI application in access to justice in Colombia was done in partnership with civil society organisations, which raised concerns regarding the first version of PretorIA. The Constitutional Court modified the project as a result, and adopted more transparent technologies. Nonetheless, persistent challenges related to privacy of information requires the AI system to continually adapt its provisions to reflect developments in the Colombian legal framework.

Justiça para todos in Brazil

In 2020, UNDP Brazil launched the project *Justiça para todos* in collaboration with the Conselho Nacional de Justiça (CNJ), an institution aiming to improve the work of the Brazilian judiciary in terms of accountability, administration and process efficiency. Under the broader objective of promoting innovation to improve access to justice, the project Justiça 4.0 includes a significant AI element. Judges' workload in Brazil is one of the highest in the world. The project is still in its initial phase. It started by completing a needs assessment of Brazil's justice sector on how to strengthen access to justice through technology. But on the longer term, it aims to develop AI solutions to improve access to justice.

AI systems' capacity to survey and process large amounts of data can help identify patterns and trends, which can be helpful in supporting judges to navigate case proceedings. AI support to judges would enable them to handle cases more efficiently which, in turn, increases their case coverage capacity. As a result of judges being able to proceed cases more rapidly, court congestion should be reduced. This may well improve citizens' access to justice, since delays in justice provision would diminish. Additionally, by fostering efficiency, the project may improve the population's perception of judiciary institutions. As the institution would gain greater legitimacy, population distrust should become a lesser barrier to access to justice. Improving justice provision and access to justice are central to SSG/R programming and AI has shown promising prospects in that area.

Justiça 4.0 is an ambitious project aiming to tackle structural obstacles hindering effective justice provision in Brazil by taking a whole-of-sector approach. The impact of the project is meant to improve transparency, access to justice and justice efficiency and build the capacity of judiciary personnel. The most important barriers to project implementation identified by the project team at this early stage is the lack of coordination with involved stakeholders. Taking a whole-of-sector approach means mobilising actors at the federal and sub-national levels that are not necessarily used to working together. Having a platform with relevant processes whereby these stakeholders can articulate their need and their vision regarding the project. Having this platform would facilitate closer cooperation on inter-institutional activities and speed up implementation.

AI and Governance

E-Services

The COVID-19 pandemic led many authorities to widen digital access to services. E-Services or E-Government offers many possibilities for improving security and justice services, such as better services accessibility and improved social accountability. With the right technology, digital records have the potential to increase transparency of security and justice sector's activities and services, helping in demystifying the security and justice space for people and civil society organisations.

Improved Transparency of procurement practices in Ukraine

With support from a coalition of civil society actor, Dozorro is a non-profit civic tech project set up to detect and prevent misuse of public funds in public procurement. The Dozorro platform uses a risk-assessment tool powered by AI to analyse the database generated by public procurement activities. Based on risk indicators, the programme is capable of red-flagging public tenders with a high-probability of violations. The identified tender is then submitted for further review to a network of 25 CSOs representing the Dozorro community. The system has been able to identify more than 20'000 violations and engaged with more than 900 000 unique users.

Not only does Dozorro facilitate civil society engagement in external oversight of public procurement, but it also fosters a culture of cooperation between different Civil Society stakeholders. This network of CSOs developed replicable modes of cooperation which should enable this community to better anchor itself as an actor in the national reform dialogue. The improved scale and political weight of the Dozorro community should translate into greater agency power which, in turn, acts as a deterrent for whomever would attempt to misallocate public funds.

To build Dozorro, 20 experts evaluated a dataset of 3 500 tenders to determine whether a tender carried risk or not and the resulting data was used to train the AI system. The AI system independently assesses the likelihood of corruption risks and flags potential risk to experts. Once the expert reviews the case and determines if there is a risk, the expert's answer is registered in the system, whether positive or negative, which allows the algorithm to keep learning and improve its accuracy.

Early attempts at automating the tender verification process were based on 35 fixed indicators such as missing signatures or documents. However, these indicators became known, and therefore vulnerable to those seeking to game the system, find loopholes and avoid detection. The development of Dozorro to use a flexible set of indicators significantly mitigates this vulnerability.

The replicability of this model in Moldova and Poland is being discussed among civil society organisations.

Electronic Voting

During the COVID-19 pandemic, numerous elections had to be postponed or cancelled. Postal voting was often advocated as an alternative solution to physical voting; however, the process remains a logistical challenge involving significant costs. Electronic voting technologies have, as a result, become an interesting alternative. The challenges of e-voting are multiple, ranging from lack of adequate technical skills; lack of voters' trust in technology; potential system attacks and viruses, to risks related to multiple voting and to the complexity of correct voter identification. Nevertheless, the potential for e-voting is expected to grow. The generational gap regarding technical skills is decreasing as internet use is growing and as we accumulate experience, the trust in the technology is also growing. Solutions such as, electronic voting IDs have already been successfully used. Ultimately, e-voting represents a valuable opportunity to guarantee the functioning of public institutions in the event of a national or global crisis.

E-voting in Burkina Faso

The Satellite Agency, SES, supported Burkina Faso in the 2015 election by installing 368 VSATs (Very Small Aperture Terminals) across the country. This enabled elections to be electronic, transparent and broadcasted live on TV. This was very successful, and last year ENABEL and SES signed a multi-year framework to deliver similar satellite-based infrastructure and services across 20 African countries[6].

AI and National Budgets

Electronic budgeting refers to the digitalisation of budgetary procedures or services to improve financial governance of security sector actors. Budgeting can be improved through the use of information and communication technologies (ICT) by digitalising budget procedures and diffusing budgetary information to the public. Electronic budgeting is a powerful tool as a cost-saving efficiency measure, as well as a step towards increased financial transparency. Automating parts of a procurement and budget processes, should in principle save staff time on auxiliary tasks, optimising processes and providing technological safeguards against corruption.

AI and Accountability

The use of AI and its technological solutions offers valuable benefits to predict and prevent crimes, violence and conflicts. However, oversight mechanism play a key role in ensuring an accountable and fair use of these technologies.

Technological Accountability

In order for AI design to be technically accountable it must guarantee system integrity, task efficacy, transparency and interoperability with other systems. The technology should only operate within the realm of its human-defined oversight framework and not overtake the relevant authorities' capacities to monitor and evaluate its functioning. In concrete terms, data used to build AI tools needs to come from certified sources and the storage and execution of algorithms have to occur in secure environments to guarantee the system's integrity and intangibility. Technical

[6] https://www.ses.com/press-release/belgian-development-agency-partners-ses-connect-foreign-aid-projects-africa-satellite

transparency can be applied through, for example, open-source code and documentation, with detailed, publicly-accessible explanation of AI systems, including functioning of services and error margins.

Legal Accountability

The legal oversight of AI solutions can be a very complex process, including public and private actors involved in digital, cyber and AI technologies and providing public service. Ensuring that AI systems respect the rule of law is essential to preserve public trust. Given that AI systems deal with sensitive and personal data, they should be in full compliance with law requirements on data protection and the processing of personal information. AI challenges legal frameworks as it can lead to “attribution confusion” since AI is neither a moral agent nor a legal person. Therefore, accountability should be based on the decision-making process to ensure that an identified individual will remain accountable for a certain course of action. When an AI system is sourced from a third-party, contracts should clearly define which organisation is liable under what circumstances.

Political Accountability

Whilst technological and legal oversight, should monitor the performance of AI applications, political oversight is necessary to ensure that the “right” technology is used for the “right” purpose, using the “necessary” data through the adequate decision-making process and in a transparent, impartial and fair manner. AI decision-making processes should remain explainable and traceable for human decision-making, which should be empowered and not restricted by the use of AI tools and services. As such, public authorities should grant a certification for the use of AI tools and independent authorities should be able to audit processing methods. A specialised commission could also regularly monitor and evaluate the use of AI in the security sector by organising planned and random audits. Such a commission could consist of members from various fields relevant for SSG/R and AI such as telecommunication, education and human rights.

Ethical Accountability

Ethical safeguards can be based on multiple foundations since they have different applications across police, defence and the justice sector. AI systems for the security sector could either emphasise the so-called ‘ethics by design approach’, meaning that processed data is reviewed by an ethic advisory board and make data and algorithms easily understandable for people. A ministerial ethics committee could further oversee the development of AI technology in the security sector and work in close collaboration with a National Ethics Advisory Committee. Another way to include ethical considerations is through the so-called ‘human rights by design approach’, which is a way of turning the UN’s Declaration of Human Rights towards the digital world. This approach puts the user at the centre, focusing on user consent and leaves several choices for the user.

In the defence sector, it is crucial that the development of AI abides to principles of international humanitarian law regarding conflicts (necessity, humanity, proportionality, distinction). The development of fully autonomous AI systems in the defence sector can violate international humanitarian law principles. Concerning the police and justice sectors, fairness and non-discrimination principles are critical in the development of AI systems.

Where To?

The crossover of SSG/R and the security sector is multifarious, and practitioners and researchers are beginning to take stock of what has already happened and where things might go. AI is changing the security landscape, the capabilities of security institutions and adding to the toolbox of reformers.

From facial recognition in conflict zones to flagging corruption risks, AI is a potent tool for state and nonstate actors alike, provoking difficult questions about benefit, autonomy, privacy and respect of human rights.

The nature of AI, which allows autonomous learning from huge data sets, makes it doubtless that security and justice sectors around the world will take up the technology to a massive degree. The speed of this uptake, and the speed of advances in AI, make the tasks of governing it well difficult. Indeed, governance frameworks are already struggling to catch up.

While the rate of change may appear dizzying, what is important to remember from a security governance viewpoint are the essential principles of people-centric security, accountability and human rights. The questions to ask therefore about the application of AI technologies are, *how can the technology make security provision more people centric? How can this technology improve accountability as well as efficiency? How can this technology be used to buttress human rights standards?* This is the essence of the SSG/R lens on AI.