U4 ANTI-CORRUPTION RESOURCE CENTRE

U4 BRIEF 2025:1

Dieter Zinnbauer

Series editor
Sofie Arjon Schütte

Artificial intelligence in anti-corruption – a timely update on AI technology

CMI CHR. MICHELSEN INSTITUTE

Corruption erodes sustainable and inclusive development. It is both a political and technical challenge. The U4 Anti-Corruption Resource Centre (U4) works to understand and counter corruption worldwide.

U4 is part of the Chr. Michelsen Institute (CMI), an independent development research institute in Norway.

www.u4.no

u4@cmi.no

Read online

www.u4.no/r/BR2501

AI has been successful in a number of 'classic' anti-corruption areas – procurement integrity, compliance, fraud detection and anti-money laundering. But AI has not gained public trust (it is difficult to explain exactly how AI works) and it has not been able to overcome challenges such as resource constraints, data quality, organisational resistance and digital divide issues. AI can also add inherited biases, and it has been responsible for erroneous outputs. However, there are some promising new uses for AI in anti-corruption work: remote sensing via satellites; and managing large-scale citizen consultations.

Main points

- Artificial intelligence (AI) has developed rapidly in the last few years. It has demonstrated impact in a number of 'classic' anti-corruption areas – procurement integrity, compliance, fraud detection and anti-money laundering.
- AI grapples with some of the same challenges that have been found to undermine the efficacy of previous generations of digital technologies in anti-corruption work. These challenges include resource constraints, organisational resistance and digital divide issues.
- AI systems come with several specific challenges that can also adversely affect deployment in the anti-corruption context. These include inherited biases, erroneous outputs, and lack of explainability – meaning the reasons why AI has arrived at a particular output cannot be fully explained.
- Emerging evidence points to several promising application areas for AI in the broader integrity environment that have so far received limited attention by anti-corruption practitioners. These areas include the use of AI to generate new insights from remote sensing via satellites, and its utility in managing large-scale citizen consultation and engagement projects.
- Donors can play a timely and important role in building targeted capacity for tailoring the latest AI systems that work as multi-purpose platforms to the fight against corruption in specific applications areas.
- Donors can also serve as important catalysts for cross-practice collaborations, and to reaffirm the focus on addressing multiple digital divide issues that hold back the realisation of many gains from AI.
- Data quality and availability are key for effective AI deployment against corruption. Donor investments in conventional accountability data (eg, asset disclosure, political finance data) and AI training data can make critical contributions.

Contents

What is artificial intelligence?

Definitions of artificial intelligence (AI) vary, but typically agree on the understanding of AI as computing systems with the following advanced capabilities that interact in a virtuous circle of continuous improvement:

1. Perceiving and extracting information from the environment
2. Processing this information and learning from it
3. Taking decisions towards a defined goal with some degree of autonomy.[1]

AI is not a new invention: it builds on a continuum of technological developments around computer algorithms, big data analytics and machine learning that have been invested in since foundational research in the 1950s. The recent growth of AI has been driven by technological breakthroughs in learning techniques, an explosion of available digital data to ingest, and rapidly growing computing power. For example, the cost of computing has been reduced by a factor of trillions since the early days of AI.[2] These systems now match, and in some instances outperform, humans in a number of intelligence benchmarks.[3]

Artificial intelligence and anti-corruption – time to check-in on new technology

Artificial intelligence (AI) has been developing in leaps and bounds over the last two decades.[4] This rapid progress has also increasingly captured the imagination of the anti-corruption community, who have long sought to harness the potential of digital technologies for anti-corruption work. These technologies have had an evolving set of thematic labels from 'e-government' and 'smart cities'[5] to 'big data',[6] 'crowd-sourced transparency',[7]'algorithmic accountability',[8] or 'integrity via blockchains'.[9]

1. Russell and Norvig 2021.
2. Leech et al. 2024.
3. Stanford University 2024.
4. Stanford University 2024.
5. Townsend 2013.
6. Berliner and Dupuy 2018.
7. Zinnbauer 2015; Grossman, Platas and Rodden 2018.
8. Kossow, Windwehr and Jenkins 2021.
9. Aarvik 2020.

Around 2020, a flurry of reports – including by U4 – began to more systematically map the implications of contemporary AI systems for anti-corruption work.[10] These important contributions had to rely mainly on early-stage experiments or hypothetical cases, as actual implementations, impact assessments and empirical evidence, were still some way off.[11]

Generative AI has experienced the fastest adoption rate of any new technology ever.

These early accounts also predate the latest transformational breakthrough in AI: the arrival of ChatGPT and other so-called generative, general purpose AI. These AI systems are capable of interacting with non-expert users and producing human-like outputs in a wide range of application areas. Generative AI has experienced the fastest adoption rate of any new technology ever and drives a fresh wave of AI investments, experimentation, and related policy conversations to new heights – and hypes. This U4 Brief provides an update on some important developments in use of AI in anti-corruption. It focuses on important emerging lessons, and points to two overlooked application areas that are particularly fertile. It also emphasises policy implications and potential engagement opportunities for donors at this stage of AI in anti-corruption deployment.[12]

The emergent learning landscape – solid progress but no magic solution yet

The expanding evidence base on AI in anti-corruption yields a number of emerging insights that often reaffirm experiences with prior generations of digital technologies. At the 'big picture' level these include:

- A proof of concept is not a failsafe ticket to sustainability or diffusion: several early examples of AI applications in anti-corruption have been discontinued or shown limited scalability. Limiting factors such as resource constraints, poor data quality, dwindling citizen engagement, rapid technical obsolescence, ill-suited regulation, or poor integration of AI solutions into existing anti-corruption infrastructures are well-known problems for many types of digitisation projects.[13] It is not surprising that they also pose obstacles for AI supported applications.[14]

10. Aarvik 2019; Coalition for Integrity 2021; Köbis et al. 2021.
11. Adam and Fazekas 2021.
12. This brief focuses on AI for anti-corruption. It does not address the use of AI for corrupt purposes, which is a growing area of concern.
13. Zinnbauer 2015; Adam and Fazekas 2021; Mattoni 2024.
14. Odilla 2023; Dreyling et al. 2024.

- A political economy and institutional lens merit more attention, when assessing technological opportunities and implementation strategies. No anti-corruption technology will succeed if the very actors tasked with deploying it are those who also benefit from corruption.[15] Teasing out the implications of AI for specific configurations of interests and power is essential for drawing up effective implementation pathways. It may be necessary to build on an expanding mix of approaches to analyse political economy contexts or institutional opportunity structures.
- Digital divide issues are becoming even more critical. Women, ethnic minorities, and other marginalised communities are underrepresented in AI research. These groups are also often mischaracterised in historical training data: they are confronted with ill-tailored applications at best, and biased, discriminatory AI outcomes at worst.[16]

Several application areas are prolific hives of experimentation and learning. Two of these are profiled below: 'classic' anti-corruption areas; and reducing corruptible discretion in decision-making.

Making good progress: AI in 'classic' anti-corruption areas

AI is making good progress in some 'classic' anti-corruption areas: procurement, compliance, collusion, and anti-money laundering (AML). And AI continues to expand capabilities in areas where digital technologies already make a sizeable impact. AI is making significant contributions in a number of application areas[17] – particularly with its ability to process and connect unstructured information at scale, and its flexible approach to unsupervised learning.

AI is making good progress in some 'classic' anti-corruption areas.

For integrity in public procurement, the use of AI can build on a long track record of digitising procurement data and processes. It also helps to advance capabilities in three main areas:

- AI makes it possible (in principle) to identify new and flexibly adjust existing red-flag indicators. This makes it more difficult for criminals to game the system. It is easier for law enforcement to catch up with new fraudulent tactics, and so

15. Adam and Fazekas 2021.
16. Odilla 2024.
17. Han et al. 2020; Raj et al. 2024; Gandhi et al. 2024.

possibly also reduce undue risk aversion in procurement.[18]

- AI enables a new generation of analysis and monitoring initiatives to consider a much broader set of data inputs – from insolvency to political connections, open ownership to open contracting vanguards, and asset disclosures to political sponsorship information – to detect more complex patterns of collusion[19]and conflicts of interests.[20]
- AI paves the way for procurement analysis at unprecedented scale and scope.

These emerging efforts can take advantage of a number of initiatives that are assembling big data repositories and large-scale red flag assessments of procurement data – for example, by compiling 17 million tenders across the European Union (EU) (DIGIWHIST)[21] or by collating data from no less than 42 national procurement systems (ProACT).[22]

AI helps to turn episodic spot-checks and select audits into more efficient, comprehensive and real-time monitoring efforts.

In the broader areas of compliance, fraud detection, and anti-money laundering, AI helps to turn episodic spot-checks and select audits into more efficient, comprehensive and real-time monitoring efforts.[23] For instance, a global beverage conglomerate consolidated more than a dozen internal enterprise resource management systems with a number of external data streams. The result was a consolidated, AI-supported supplier vetting function that reduced costs by more than 90%.[24]

In Peru, investigators use AI to screen a growing volume of reported suspicious financial transactions. The process is much more efficient, doubling the rate of cases that could be referred to prosecutorial authorities.[25] Similarly, a global bank has cut the incidence of false positives in fraud reporting by 75%.[26] Another bank doubled its

18. Menke et al. 2024.
19. Ghedini Ralha and Sarmento Silva 2012; García Rodríguez et al. 2022.
20. Mazrekaj, Titl and Schiltz 2021; Cocciolo et al. 2024.
21. Petheram 2019.
22. Cocciolo et al. 2024.
23. FATF 2021.
24. Coalition for Integrity 2021.
25. GIZ 2024.
26. The Global Treasurer 2024.

detection rate of confirmed bad transactions and cut transaction processing time from more than a month to just a few days.[27]

Mixed success: AI in administering social benefits or reducing corruption-prone discretion in other areas

It is clear that AI can help administer standard cases and free up human resources for more complex issues. However, high hopes that AI could play a major role in cutting corruption at public service level by reducing corruptible discretion in decision-making have so far remained unrealised for several reasons, including:

- Eliminating discretion remains an ill-conceived ambition: retaining some discretion is a desirable principle in administrative decision-making to arrive at fair outcomes that are tailored to individual circumstances.
- AI systems were found to produce many incorrect decisions in several high-profile implementations – from denying unemployment benefits to thousands of rightful claimants in Michigan, causing dire economic distress,[28]to wrongfully withdrawing social benefits in Serbia,[29] or child benefits allowances from Dutch parents – a scandal that ultimately led the government to resign.[30]

These shortcomings exemplify some fundamental problems with AI in this context. AI operates largely as a 'black box', partly due to technical complexity, partly due to proprietary ownership of AI models and data. Because AI decisions cannot be fully explained, this violates a basic principle of administrative justice. Biases in training data and the persistent, significant tendency to make things up further undermine accuracy and fairness. ('Hallucination' is the technical term used to describe an AI-generated response that contains misleading information presented as fact.) Even AI systems that specialise in legal issues are found to 'hallucinate' in up to 30% of cases – for example, referencing legal clauses that do not exist.[31]

A 'human in the loop' as ultimate decision-maker is a prerequisite for achieving just and accurate outcomes.

All these issues suggest that AI can provide decision support, but that a 'human in the loop' as ultimate decision-maker is a prerequisite for achieving just and accurate

27. Katkov 2023. See also Google 2024.
28. Ceva and Jiménez 2022.
29. Cvetinovic 2023.
30. Peeters and Widlak 2023.
31. Magesh et al. 2024.

outcomes – and also to establish clear lines of accountability when things go wrong. As much as discretion might invite corruption, an over-reliance on AI at this stage is not desirable and produces materially bad outcomes. Also, AI systems can potentially be tweaked to produce particularistic outcomes (those that favour a particular population or group).

These shortcomings are not insurmountable, and many efforts are underway to mitigate them – from devising assessment frameworks for model transparency,[32] to ethics benchmarks,[33] and novel approaches for explainability.[34] However, given the rapid development of AI, many of these initiatives are in catch-up mode, and issues continue to persist or reappear in different forms. Also, in the quest for proprietary advantage, many cutting-edge AI models do not include sufficient public disclosures.[35]

New anti-corruption spaces for AI

So far, much attention in policy and practice has focused on 'classic' anti-corruption applications – from compliance to social accountability. But there are also a number of secondary application areas where emergent evidence shows a growing impact and future high potential for use of AI in anti-corruption. Two areas are profiled below: remote sensing; and inclusive participation.

Remote sensing with anti-corruption momentum

Over the last ten years, AI systems have developed their ability to extract information from images and recognise complex patterns. In parallel, we have seen a rapid expansion of available satellite imagery and earth observation data. The number of satellites in space has tripled within the last five years to reach 10,000.[36] These satellites now cover Global North and Global South countries and produce reams of up-to-date data at an unprecedented level of geographic detail. This opens up novel opportunities to harness AI for making progress on tackling a number of illicit activities that are particularly consequential, deeply interwoven with corruption, and particularly prone to impunity.

32. Bommasani et al. 2023.
33. Stanford University 2024.
34. Molnar 2020.
35. Leech et al. 2024.
36. Rainbow 2024.

A new cohort of alert systems use the predictive power of AI to identify early warning signs of deforestation.

For example, satellite-based monitoring of illicit logging and deforestation has improved significantly, enabling sub-weekly updates at a spatial detail level of 10 to 50 metres.[37] Related monitoring initiatives are now able to map deforestation in much more detail, and also track specific activities that drive it, including illegal gold mining in the Amazon,[38] land grabs in Ghana,[39] or illegal coal mining in China.[40] A new cohort of alert systems (currently in experimental stages) also use the predictive power of AI to identify early warning signs of deforestation – for example, road buildings or adjacent infrastructures such as mills or silos[41] – and can cut response times significantly.

These initiatives are already bearing fruit. Forest Foresight – a collaboration between academia and a non-governmental organisation – claims to be able to predict illicit deforestation months before it occurs. In a trial implementation in Gabon, it helped park rangers carry out 34 enforcement actions and stopped an illicit gold mine.[42]

Combined with other datasets, some of the underlying corruption dynamics can be much more effectively documented. Deforestation in Brazilian communities was found to spike by another 8% to 10% in election years when an incumbent mayor runs for re-election, pointing to the use of forbearance as a political tool (where those in authority withhold certain sanctions to gain votes).[43] Meanwhile, enforcement action by federal-level agencies during the 2019 to 2022 presidential term was identified to be minuscule, with less than 1% of incidences receiving a response.[44]

Similarly, AI-supported remote sensing analysis is being used to better understand, track, and create the prerequisites for more accountability in a wide range of corruption-prone application areas, such as :

- Tripling the productivity of rangers to detect snares in Cambodia[45]

37. Reiche et al. 2024.
38. Labbe 2021.
39. Hausermann et al. 2018.
40. Xia et al. 2023.
41. Hillsdon 2024.
42. World Wildlife Fund 2023.
43. Pailler 2018.
44. Coelho-Junior et al. 2022.
45. Xu et al. 2019.

- Assigning responsibility for remote oil spills in the Mediterranean[46]
- Revealing fake suppliers in procurement in Brazil[47]
- Illegal bitcoin miners in Iran[48]
- Illicit fishing activities around the world[49]
- Ethnic patronage in Kenyan informal settlements[50]
- Bureaucratic incentives for harmful crop-burning in India.[51]

AI has also helped to detect and locate (at very detailed level): misreporting of economic activity in China[52] and of methane emissions in the USA;[53] and uneven, politically motivated government response to disasters in Brazil, Indonesia, Mexico, and South Africa.[54]

Three characteristics make this type of AI-enabled remote-sensing particularly interesting from an anti-corruption perspective.

Concerns remain about the availability and affordability of particularly high-quality satellite data. However, at least three characteristics make this type of AI-enabled remote-sensing particularly interesting from an anti-corruption perspective:

- It can bring transparency to areas and issues that are too remote or too tightly controlled by corrupt power brokers to be effectively monitored.
- It often makes this data available to civil society stakeholders and researchers outside the reach of repressive governments.
- It operates at scale, at near real-time frequency, and at a level of geographic detail that helps directly assign responsibility – for example, by pinpointing a specific polluting source, deviant actor, or local council.

46. Skytruth 2024.
47. Wacker, Ferreira and Ladeira 2018.
48. Amiri and Askari 2022.
49. Paolo et al. 2024.
50. Marx, Stoker and Suri 2019.
51. Dipoppa and Gulzar 2024.
52. Baccini, Hu and Li 2024.
53. Nesser et al. 2024.
54. Ahrenshop et al. 2024.

Also, remote sensing via satellites comes with a global geographic scope that does not leave out the Global South which is typically underrepresented in many data collection efforts.

Protecting against 'policy capture', and promoting inclusive participation

Anti-corruption and integrity systems often seek to counter risks of policy capture by vested interests through various types of public consultation, participation, and feedback systems. However, in many contexts, such mechanisms have become almost impossible to manage. Electronic submission formats have increased the number of feedback submissions and comments. Some are mass-organised, and mechanisms are increasingly targets of attempted manipulation by special interests or propagandistic actors.[55]

For example:

- The US Consumer Financial Protection Bureau receives more than 1 million comments annually.
- The public consultation for a new Chilean constitution drew more than 280,000 individual comments that had to be hand-coded in a very short time frame.
- The EU consultation on the future of Europe attracted more than 50,000 participants and 17,000 ideas that were summarised by consulting firms into lengthy reports.

Such 'mega-consultations' typically face two principal problems depending on their main objectives:

- The 'haystack challenge' to find a small number of comments that contain very high informational value among a flood of low-quality submissions.
- The 'forest challenge' to identify, consolidate and numerically weigh similar concerns and opinions for an overall overview of where public sentiments are.[56]

AI's ability to categorise and summarise very high volumes of natural language text offers ways to address these problems and make processes meaningful and effective. For example, the Consul Democracy software platform is used by more than 300 cities and organisations around the world to help with these tasks.[57] Its early iteration, pioneered in Madrid, attracted more than 26,000 proposals. Without the use of AI, time-consuming summarisation and consolidation of information meant

55. Bagley 2019; Carpenter et al. 2022.
56. Berliner 2023.
57. Ábrám 2024.

that only two proposals reached the threshold to be considered by the city council. A new, AI-supported version of the platform allows for much better search and consolidation of similar results.[58] Overall, such AI-powered consultation summary systems are becoming better. AI systems are also demonstrating their usefulness in helping to moderate deliberative processes or negotiations that have previously been bogged down by polarisation.[59] Yet AI continues to face problems, particularly with regard to non-English language applications.[60]

AI systems are also demonstrating their usefulness in helping to moderate deliberative processes.

Nevertheless, this bodes well for AI helping in very specific applications to tackle some important challenges that affect participatory and deliberative mechanisms. AI can also help to ensure that the vital pillars for inclusive governance are not hijacked by those with special interests or overwhelmed by the scale of inputs.

Implications for anti-corruption donors and practitioners

With AI in anti-corruption moving increasingly from proof of concept and experimentation to a more mature, and pragmatic deployment stage, there are a number of implications for donors and policymakers.

A reaffirmed need to address digital divide issues

High costs and illiteracy can prevent some groups from accessing and benefiting equitably from information and communication technologies. This includes across country, gender, ethnic, or socio-economic lines. These disparities are further amplified by AI, leading to a cumulative digital divide. Unfortunately, this inequity can be easily obscured due to the appearance of neutrality that surrounds AI and its statistical approaches.[61]

Equally problematic is the fact that only an estimated 22% of AI professionals are women.[62] Also, marginalised communities have low representation in the digital realm, or else they can have a disproportionate high visibility in relation to specific adverse events (eg, more prevalence in crime statistics due to more police attention).

58. Arana-Catania et al. 2021.
59. Tessler et al. 2024.
60. Romberg and Escher 2024.
61. Stone 2024.
62. Leech et al. 2024.

These distortions mean that less high-quality training data is available for these particular groups. This can lead to AI systems that are likely to produce more erroneous, often biased results. For example, AI-enabled decision-support systems in hiring personnel might reproduce gender disparities when relying on legacy data that is skewed towards hiring and promoting male candidates.[63] On the supply side, AI systems are highly concentrated in and controlled by a small number of commercial players in the Global North,[64] making it unlikely that the most advanced models speak directly to the needs of the Global South. Therefore, targeted donor support and mobilisation of investments and partnerships to bridge critical dimensions of the digital divide should rank high on the digital policy agenda. These initiatives can be rooted in the Global Digital Compact, which the global community committed itself to in mid-2024.[65]

Curate accountability data and communities

AI can serve as an integrative tool that can help extract insights from disparate and unharmonised information repositories at scale. In Armenia, for example AI is being deployed to more effectively scrutinise asset disclosures by officials.[66] In Czechia, AI helps identify long chains of political connections with finance and open ownership data.[67]

AI can help to connect disparate disclosure, transparency, and open data initiatives.

With these functionalities, AI can help to connect disparate disclosure, transparency, and open data initiatives. The technology can take the combined value of these data pools to a new level for anti-corruption analysis and monitoring. Making this possible requires cross-practice networks, which is what donors can help build. This also requires continued investments in advocacy for transparency initiatives, and maintenance of high-quality, unbiased, open government and open data repositories whose utility will be significantly amplified by AI. Where concerns about breaching regulatory provisions might delay such investments, public donors could consider helping establish experimental applications, so-called 'sandboxes', where such interlinked applications can be tested under proper scrutiny.[68]

63. Raghavan et al. 2020.
64. Stanford University 2024; Leech et al. 2024.
65. United Nations 2024.
66. Izdebski et al. 2023; Eurasianet 2023.
67. Mazrekaj, Titl and Schiltz 2021.
68. European Union 2024.

Invest in targeted training data, open ownership models, and broader data collection

Limited availability of unbiased training data is one of the major constraints in fully harnessing the potential of AI, including its application in anti-corruption. For instance, in remote sensing, an abundance of monitoring data contrasts with a dearth of ground-validated training data.[69] Targeted support to build up specific, openly accessible training datasets in close consultation with the professional anti-corruption community might help to unlock this potential.

Build resources and infrastructures for challenging unfair AI outcomes

Given the scale of AI systems' operations, it is inevitable that, even in optimal conditions, AI will produce a large absolute number of false negatives in anti-corruption screening efforts – that is, citizens erroneously accused of fraud, debanked, denied social benefits they are entitled to, and so on. Disadvantaged groups are more likely to be affected and less likely to muster the resources for contesting such decisions. Buffering the downside of AI in these areas will mean supporting affected individuals by building practical capabilities, helping them file effective complaints and, if needed, launching court cases. This can most productively be complemented by investing in analytical tools and assessment frameworks to help identify when and how AI fails to deliver, and produces biased and otherwise erroneous outcomes.

Build capacity for AI in the broader anti-corruption community

The latest conversational AI tools come with the appearance of easy accessibility and use, in a Google-search type of way. But this is treacherous. Supporting capacity-building for AI in anti-corruption is imperative for several reasons.

- In high-risk anti-corruption environments, a prerequisite of responsible AI use is to understand its many limitations, from biases to 'hallucinations'.
- A major challenge is how to most effectively query conversational AI systems, as results are found to vary widely with the quality of the query. This has given rise to a dedicated expert community of so-called 'prompt engineers'.
- The next step is using existing AI systems as a platform to customise applications for anti-corruption purposes. This requires technical AI expertise in public administrations and compliance departments. On the civil society side, we need good governance and robust 'watchdog' non-governmental organisations.

69. Burke et al. 2021.

On the horizon are opportunities to train open-source AI systems with curated anti-corruption data to build more effective mission-centred AI systems. This will require advanced development skills and pragmatic, hype-resistant awareness of AI prospects at management level. The future also calls for new collaborative efforts as resources might have to be pooled and application benefits spread beyond individual organisations.

Donors can play important roles in supporting the next phase of AI against corruption.

Donors can play important roles in supporting the required technical capacity, management awareness and collaborative spirit required for entering the next phase of AI against corruption. They can help to retain a pragmatic perspective that cuts through any hyperbole to diligently examine where and how AI can be most effectively integrated into integrity systems. Donors can also identify specific applications where the costs (still) outweigh the benefits, or where simpler technological solutions are available.

References

Aarvik, P. 2019. Artificial intelligence – a promising anti-corruption tool in development settings? U4 Report. U4 Anti-Corruption Resource Center, Chr. Michelsen Institute.

Aarvik, P. 2020. Blockchain as an anti-corruption tool. Case examples and introduction to the technology. U4 Issue. U4 Anti-Corruption Resource Center, Chr. Michelsen Institute.

Ábrám, E. 2024. AI: Supporting citizen engagement in the modern era. Transatlantic Perspectives.

Adam, I. and Fazekas, M. 2021. Are emerging technologies helping win the fight against corruption? A review of the state of evidence. Information Economics and Policy 57:100950.

Ahrenshop, M. et al. 2024. Unequal disasters and uneven response. Conference presentation: The mass politics of climate change, Philadelphia, 7 September.

Amiri, M. and Askari, H. 2022. Illegal miner detection based on pattern mining: A practical approach. Journal of Computing and Security 9 (2):1–10.

Arana-Catania, M. et al. 2021. Citizen participation and machine learning for a better democracy. Digital Government: Research and Practice.

Baccini, L., Hu, Y. and Li, B.G. 2024. Detecting corruption from outer space. Applied Economics Letters 0 (0):1–6.

Bagley, N. 2019. The procedure fetish. Michigan Law Review 118 (3):345–401.

Berliner, D. 2023. Information processing in participatory governance. Center for Open Science.

Berliner, D. and Dupuy, K. 2018. The promise and perils of data for anti-corruption efforts in international development work. U4 Brief. U4 Anti-Corruption Resource Center, Chr. Michelsen Institute.

Bommasani, R. et al. 2023. The foundation model transparency index. arXiv:2310.12941

Burke, M. et al. 2021. Using satellite imagery to understand and promote sustainable development. Science. 2001 Mar 19;371 (6535).

Carpenter, D.P. et al. 2022. Inequality in administrative democracy: Methods and evidence from financial rulemaking. Mossavar-Rahmani Center for Business and Government, Harvard Kennedy School.

Ceva, E. and Jiménez, M.C. 2022. Automating anticorruption? Ethics and Information Technology 24 (4):48.

Coalition for Integrity (C4I). 2021. Using machine-learning for anti-corruption risk-and compliance.

Cocciolo, S. et al. 2024. Power of global data and analytics in strengthening the accountability and performance of public procurement. World Bank Blogs. 28 March 2024.

Coelho-Junior, M.G. et al. 2022. Unmasking the impunity of illegal deforestation in the Brazilian Amazon: A call for enforcement and accountability. Environmental Research Letters 17 (4):041001.

Cvetinovic, A.T. 2023. Doomed by algorithm: Serbia's 'social card' leaves society's weakest exposed. Balkan Insight.

Dipoppa, G. and Gulzar, S. 2024. Bureaucrat incentives reduce crop burning and child mortality in South Asia. Nature 634, 1125–1131.

Dreyling, R. et al. 2024. Navigating the AI maze: Lessons from Estonia's Bürokratt on Public Sector AI Digital Transformation. SSRN Scholarly Paper.

Eurasianet. 2023. Perspectives: Artificial intelligence boosts anti-corruption efforts in Armenia. 12 May 2023.

European Union. 2024. Regulation (EU) 2024/1689 (EU AI Act).

FATF (Financial Action Task Force). 2021. Opportunities and challenges of new technologies for AML/CFT.

Gandhi, H. et al. 2024. Navigating the complexity of money laundering: Anti–money laundering advancements with AI/ML Insights. International Journal on Smart Sensing and Intelligent Systems 17 (1):20240024.

García Rodríguez, M.J. et al. 2022. Collusion detection in public procurement auctions with machine learning algorithms. Automation in Construction 133:104047.

Ghedini Ralha, C. and Sarmento Silva, C.V. 2012. A multi-agent data mining system for cartel detection in Brazilian government procurement. Expert Systems with Applications 39 (14):11642–56.

GIZ. 2024. Inspector AI tackles money laundering. 26 March 2024.

Google. 2024. Anti Money Laundering AI.

Grossman, G., Platas, M.R. and Rodden, J. 2018. Crowdsourcing accountability: ICT for service delivery. World Development 112:74–87.

Han, J. et al. 2020. Artificial intelligence for anti-money laundering: A review and extension. Digital Finance 2 (3):211–39.

Hausermann, H. et al. 2018. Land-grabbing, land-use transformation and social differentiation: Deconstructing "small-scale" in Ghana's recent gold rush. World Development 108:103–14.

Hillsdon, M. 2024. From forest-listening to advanced remote sensing, can AI turn the tide on deforestation? Ethical Corporation Magazine. Reuters, 16 January 2024.

Izdebski, K. et al. 2023. The digitalization of democracy. National Endowment for Democracy.

Katkov, N. 2023. HSBC: Cloud-based financial crime detection at scale. CELENT.

Köbis, N., Starke, C. and Rahwan. I. 2021. Artificial intelligence as an anti-corruption tool (AI-Act) - Potentials and pitfalls for top-down and bottom-up approaches. arXiv.Org.

Kossow, N., Windwehr, S. and Jenkins, M. 2021. Algorithmic transparency and accountability. Transparency International Anti-Corruption Helpdesk Answer. Transparency International.

Labbe, N. 2021. Detecting illegal gold mining sites in the Amazon forest: Using deep learning to classify satellites images. KTH.

Leech, G. et al. 2024. Ten hard problems in artificial intelligence we must get right. arXiv: 2402.04464.

Magesh, V. et al. 2024. Hallucination-free? Assessing the reliability of leading AI legal research tools. arXiv:2405.20362.

Marx, B., Stoker, T.M. and Suri, T. 2019. There is no free house: Ethnic patronage in a Kenyan slum. American Economic Journal: Applied Economics 11 (4):36–70.

Mattoni, A., ed. 2024. Digital media and grassroots anti-corruption: contexts, platforms and data of anti-corruption technologies worldwide. Edward Elgar Publishing.

Mazrekaj, D., Titl, V. and Schiltz, F. 2021. Identifying politically connected firms: A machine learning approach. SSRN Scholarly Paper.

Menke, W., Gomes, R. and Xavier, F. 2024. Impacts of AI-based anti-corruption audits on risk aversion in decision-making: A case study of the Brazilian ALICE tool. Global Public Policy and Governance 4 (3):273–86.

Molnar, C. 2020. Interpretable machine learning. (Self-published) Lulu.com.

Nesser, H. et al. 2024. High-resolution US methane emissions inferred from an inversion of 2019 TROPOMI satellite data: Contributions from individual states, urban areas, and landfills. Atmospheric Chemistry and Physics 24 (8):5069–91.

Odilla, F. 2023. Bots against corruption: Exploring the benefits and limitations of AI-based anti-corruption technology. Crime, Law and Social Change 80 (4):353–96.

Odilla, F. 2024. Unfairness in AI anti-corruption tools: Main drivers and consequences. Minds and Machines 34 (3):28.

Pailler, S. 2018. Re-election incentives and deforestation cycles in the Brazilian Amazon. Journal of Environmental Economics and Management 88:345–65.

Paolo, F.S. et al. 2024. Satellite mapping reveals extensive industrial activity at sea. Nature 625 (7993):85–91.

Peeters, R. and Widlak, A.C. 2023. Administrative exclusion in the infrastructure-level bureaucracy: The case of the Dutch daycare benefit scandal. Public Administration Review 83 (4):863–77.

Petheram, A. 2019. The next generation of anti-corruption tools: Big data, open data & artificial intelligence. Oxford Insights.

Raghavan, M. et al. 2020. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 469–81. Barcelona Spain: ACM.

Rainbow, J. 2024. SIA reports more record growth for the global commercial satellite industry. SpaceNews. 13 June 2024.

Raj, M. et al. 2024. The use of artificial intelligence in anti-money laundering (AML). In 2024 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL), 272–7.

Reiche, J. et al. 2024. Integrating satellite-based forest disturbance alerts improves detection timeliness and confidence. Environmental Research Letters 19 (5):054011.

Romberg, J. and Escher, T. 2024. Making sense of citizens' input through artificial intelligence: A review of methods for computational text analysis to support the evaluation of contributions in public participation. Digital Government: Research and Practice 5 (1):1–30.

Russell, S.J. and Norvig, P. 2021. *Artificial Intelligence: A Modern Approach*. Fourth Edition. Pearson Series in Artificial Intelligence. Pearson.

SkyTruth. 2024. Mediterranean sea chronic oil pollution analysis: July 2020–January 2024.

Stanford University. 2024. Artificial Intelligence Index report 2024.

Stone, E. 2024. AI shifts the goalposts of digital inclusion. Joseph Rowntree Foundation.

Tessler, M.H. et al. 2024. AI can help humans find common ground in democratic deliberation. Science 386 (6719):eadq2852.

The Global Treasurer. 2024. How AI is transforming customer service, security, and financial management in banks.

Townsend, A.M. 2013. *Smart cities: Big data, civic hackers, and the quest for a new utopia*. W.W. Norton & Company.

United Nations. 2024. Global Digital Compact.

Wacker, J., Ferreira, R.P. and Ladeira, M. 2018. Detecting fake suppliers using deep image features. In 2018 7th Brazilian Conference on Intelligent Systems (BRACIS), 224–9. Sao Paulo: IEEE.

World Wildlife Fund. 2023. Could AI help stop deforestation before it starts?

Xia, Y. et al. 2023. Combined PS-InSAR technology and high-resolution optical remote sensing for identifying illegal underground mining in the suburb of Yangquan City, Shanxi Province, China.Remote Sensing 15 (14):3565.

Xu, L. et al. 2019. Stay ahead of poachers: Illegal wildlife poaching prediction and patrol planning under uncertainty with field test evaluations. arXiv: 1903.06669.

Zinnbauer, D. 2015. Crowdsourced corruption reporting: What petrified forests, street music, bath towels, and the taxman can tell us about the prospects for its future. Policy & Internet 7 (1):1–24.

About the author

Dieter Zinnbauer

Dr. Dieter Zinnbauer is a research fellow at Copenhagen Business School. He has more than 25 years of experience in governance analysis, training, and strategy advice for intergovernmental and nongovernmental organisations, think tanks, and academic institutions.

Acknowledgements

The author and U4 thank Claudia Saenen, Fernanda Odilla, Isabelle Adam, Kim Schultze, and Matthias Luetjen for their helpful comments on an earlier draft of this paper.

Keywords

anti-corruption interventions – anti-corruption tools – artificial intelligence – big data – citizen engagement – digital technology – innovation

How to cite

Zinnbauer, D.; (2025) Artificial intelligence in anti-corruption – a timely update on AI technology. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute (U4 Brief 2025:1)

Publication
First published 7 January 2025

U4 partner agencies

German Corporation for International Cooperation – GIZ

German Federal Ministry for Economic Cooperation and Development – BMZ

Global Affairs Canada

Ministry for Foreign Affairs of Finland

Ministry of Foreign Affairs of Denmark / Danish International Development Assistance – Danida

Norwegian Agency for Development Cooperation – Norad

Swedish International Development Cooperation Agency – Sida

Swiss Agency for Development and Cooperation – SDC

UK Aid – Foreign, Commonwealth & Development Office

Corruption erodes sustainable and inclusive development. It is both a political and technical challenge. The U4 Anti-Corruption Resource Centre (U4) works to understand and counter corruption worldwide.

U4 is part of the Chr. Michelsen Institute (CMI), an independent development research institute in Norway.

www.u4.no

u4@cmi.no