Artificial intelligence as a weapon to fight corruption: Civil society actors on the benefits and risks of existing bottom-up approaches

Forjan, J.; Köbis, N. ; Starke, C.

DOI
10.4337/9781802202106.00020

Publication date
2024

Document Version
Final published version

Published in
Digital Media and Grassroots Anti-Corruption

License
CC BY-NC-ND

11. Artificial intelligence as a weapon to fight corruption: Civil society actors on the benefits and risks of existing bottom-up approaches

Julia Forjan, Nils Köbis, and Christopher Starke

INTRODUCTION

Corruption is a global, timeless, multifaceted phenomenon with severe consequences for society, politics, and the economy (Mungiu-Pippidi & Heywood, 2020; Rothstein, 2011). Against the backdrop of an increasingly digital society and the associated availability of big data, information and communication technologies (ICTs) to fight corruption are surging. These tools can provide new opportunities to hold public officials accountable for their actions (Mattoni, 2021). Consequently, a growing number of studies are investigating the potential of ICTs to combat corruption (Adam & Fazekas, 2021; Kossow & Dykes, 2018).

While classical ICTs are static, more recently, the use of the more dynamic and autonomous technologies summarised under the umbrella term Artificial Intelligence (AI) has garnered interest within anti-corruption communities. AI is defined here as 'systems that display intelligent behaviour by analysing their environment and taking actions (with some degree of autonomy) to achieve specific goals' (Ala-Pietilä et al., 2019). Such technologies are increasingly coming to the forefront of corruption research (Aarvik, 2019; Adam & Fazekas, 2021; Köbis et al., 2022).

Recently, Microsoft announced its AI Technology Solutions project aiming to employ technology to counter corruption. Shortly after, it announced a cooperation with the Nigerian government 'to apply technologies like artificial intelligence and machine learning to help identify potential risks, highlight them, and reduce corruption' (Microsoft, 2021). Not only have governments in liaison with private companies recognised AI as a promising tool to fight corruption., also, non-governmental organisations (NGO) and donor organisations, like the World Bank, have begun to draw on AI technology for this purpose (Global Witness, 2021; World Bank, 2020). Consultancies are already heralding the technology as 'the next frontier in anti-corruption' (Petheram, 2018).

But what makes AI a promising anti-corruption technology? The most crucial advantage of AI-based anti-corruption technologies (AI-ACTs) lies in how they function. First, AI-ACTs differ from classical ICTs in that they can autonomously execute tasks previously solely performed by humans (Domingos, 2012; Rahwan et al., 2019). Second, due to their large computing capacity, AI-ACTs are able to analyse large volumes of data, including Big Data and large-scale datasets, in the shortest possible time and thereby uncover suspicious patterns (Ponti et al., 2021). One example is *Arachne*, a fraud risk-scoring tool launched by the European Commission, which aims to 'detect and prevent potential irregularities in projects or contracts' (European Commission, 2021). Third, AI-ACTs operate independently of external factors that would influence their effectiveness and impartiality. Unlike humans, they can neither be pressured by time constraints nor supervisors.

Moreover, Big Data becomes increasingly available (Petheram et al., 2019). With ever more data available for analysis purposes, there has been a surge of research interest in AI for anti-corruption in recent years. However, empirical studies on the use of AI-ACTs are still scarce. To this point, they have been mainly applied in related areas, such as crime (Li & Juhola, 2015), credit risk assessment (Swiderski et al., 2012; Mhlanga, 2021), and financial fraud detection (Olszewski, 2014; Choi & Lee, 2018).

In one of the first empirical studies, López-Iturriaga and Pastor Sanz (2018) developed a model of neural networks to predict public corruption in Spanish provinces. The authors identified multiple economic and political factors that increase public corruption, such as taxation of real estate, economic growth, or the same political party remaining in power for a long period. Furthermore, the results demonstrate that the model could forecast corruption in some provinces up to three years before corruption cases were observed (López-Iturriaga & Sanz, 2018).

Another study, which aimed at determining predictors of corruption, was conducted by Lima and Delen (2019). In their cross-sectional study, the authors applied diverse machine learning techniques to identify predictors for Corruption Perception Indices across 132 countries. Based on their models, the authors identify factors such as government integrity, judicial effectiveness, and education index to influence perceptions of corruption.

On a more local level, Ash et al. (2020) and de Blasio et al. (2020) employed machine learning techniques to predict local-government corruption in Brazilian and Italian municipalities, respectively. Another study by Colliri and Zhao (2019) applied a network-based approach to analyse bills-voting data comprising the votes of Brazilian members of Congress. They argue that bills-voting networks can be used to identify politicians who are involved in corruption or other financial crimes. Mazrekaj et al. (2021) focused on identifying political connections that could represent significant conflicts of interest. They used machine learning techniques to predict politically connected firms by constructing a novel firm population dataset in the Czech Republic.

The abovementioned studies illustrate that AI-ACTs, mainly machine learning techniques, can recognise and predict patterns indicative of various types of corruption in datasets. This autonomous ability makes it possible to take preventive rather than reactive anti-corruption measures. Besides overview articles summarising AI as an anti-corruption tool (Aarvik, 2019), there are theoretical works that have started to discuss the implications of these tools from a socio-technical perspective (Köbis et al., 2022). Building on this theoretical groundwork, this study adds empirical flesh to the bones by assessing the benefits and risks of AI to curb corruption. This is made through the lens of civil society actors, who lead existing pilot projects using such technologies. The assessment of civil society actors seems particularly relevant, considering that state bodies are often unwilling or unable to take action against corruption themselves. Especially in countries where corruption is systemic, civil society actors play a crucial role as watchdogs (Mattoni, 2021). Hence, we conduct qualitative interviews with relevant stakeholders who employ AI-ACTs in the fight against corruption.

METHODOLOGY

Since this study aims to explore the benefits and risks of using AI to fight corruption, we use a qualitative approach to shed light on this complex topic. While the application of AI systems in anti-corruption efforts is surging, empirical research on its benefits and risks is lagging behind (for an exception, see Odilla, 2023). In particular, a systematic investigation of stakeholders directly involved in the implementation of AI-ACTs is lacking. To fill this gap, we conducted interviews with civil society actors as a method of inquiry.

Sampling

The selection of civil society actors followed a successive strategy using the method of purposeful sampling. This approach selects specific cases based on relevant characteristics (Palinkas et al., 2015). Thus, the experts were not selected under aspects of representativeness, but on the basis of their professional expertise and experience within the field of research interest (Littig, 2009). Since this study investigates bottom-up initiatives that use AI-ACTs, the internal expertise of the civil society actors is of particular interest. Consequently, those stakeholders were selected for the interviews with sufficient operational knowledge. They acquired this either by developing and/or participating in an AI-ACT project.

In addition, we further used snowball sampling. This method is a mixture of deliberate and random selection (Palinkas et al., 2015). At the end of the interview, the expert is asked to name other experts in the field. This procedure is particularly suitable for the exploration of novel subjects of investigation in which the population of experts is either unknown or difficult to access (Wroblewski & Leitner, 2009). At the time of data collection (August 2019–July 2020), only a few research projects and pilot studies existed that employed AI systems with the concrete goal of fighting corruption. In addition, we balanced the experts according to their field of activity in order to realistically depict the potentials and limitations of cutting-edge AI technologies. In this way, we could obtain diverse opinions on the use of AI in anti-corruption. Applying this strategy also means that our sample of experts is not representative but rather based on the experts' professional predisposition and experience (Littig, 2009). Therefore, this study includes experts who have been or are still actively involved in the development and implementation of AI-ACTs.

Since we assured anonymity to the interviewees, neither their names nor their projects' names are mentioned in this chapter. However, to make the expertise of the civil society actors comprehensible, information about their academic and professional background and their function within the project is disclosed. All of the interviewed experts are employed by civil society organisations, except for one expert who worked for a civil society organisation until they transferred to a governmental agency. The interviewees had diverse academic backgrounds, including software engineering, computer science, political science, economics and econometrics. One expert is the co-founder of a South American technology project which employs AI to track public expenditures. Three other experts were involved in public procurement projects which use AI-based tools to uncover corruption. One of them is the co-founder and director of a think tank based in Eastern Europe. They were directly involved in developing and maintaining a tool that scans and pre-filters risky public procurements. Another interviewee works for another international NGO, also based in Eastern Europe, where they monitor and coordinate public procurement tools. The other interviewee works for a think tank where they conduct research and develop tools to detect corruption in public procurement. Another expert works for a governmental agency in South America where they develop and implement an AI-based tool that assesses the corruption risk of individual government officials.

Interview Guideline

We designed an interview guide to ensure the comparability of the data on the one hand and to maintain the necessary openness for individual narratives on the other. The guide starts off with a short briefing that explains the research project and its objectives to the interviewee. In a subsequent step, the interviewee provided informed consent to record the interview.

The actual interview commenced by asking the opening question to establish a narrative conversation: 'To start off, I would like you to briefly describe how you came to work on AI and anti-corruption in general and specifically on your project'. The main part of the interview was structured in four blocks:

1. definition of the key terms, AI and corruption,
2. description of the interviewee's project,
3. potentials & limitations of using AI for anti-corruption,
4. outlook of additional fields in which AI could be applied.

The block sequence served as a broad guideline instead of a strict rule to avoid abrupt topic changes that interrupted the interview's flow (Miles & Gilbert, 2005). For this chapter, we focus on the third block which we further divided into four sub-dimensions: (a) technological, (b) human resources, (c) legal and (d) social/ethical. Finally, we asked the experts, according to the snowball procedure, to refer to other projects or experts.

Data Collection and Data Analysis

The interviews were conducted via online video tools, such as Skype or Google Hangout. In total, we interviewed five experts working on projects in four different countries in Europe and South America. The average length of the interviews was 78.5 minutes, with 127 minutes being the longest and 58 minutes being the shortest. All of the interviews were digitally recorded and transcribed in order to analyse the text transcripts afterward. For the analysis of the qualitative data, we used the software Atlas.ti.

We performed a qualitative content analysis (Mayring, 2015) of the data. Accordingly, the expert interviews were evaluated in three steps: (1) summary, (2) explication, and (3) structuring. First, we summarised relevant information that contributed to answering the research question to reduce the text material. Second, we added explanatory information to ensure the comprehensibility of the respective passage. Third, we derived seven categories from the text and assigned them to the previously selected text passages.

RESULTS AND DISCUSSION

The semi-structured interviews revealed that civil society actors see three key potentials (efficient use of human and financial resources, enhancing civic engagement, increasing transparency) and three fundamental limitations (insufficient quality and quantity of data, legal restrictions, technological challenges) of using AI in anti-corruption efforts.

Key Potentials

Efficient use of human and financial resources

All interviewees reported the time- and cost-saving function of AI-ACTs. One civil society actor reported that the government in their country does not have the human resources to control all expenditures by the members of parliament:

> The department of the congress is receiving about 2000 receipts per day. They don't have time to check all of them. Not because they don't know the law. They don't have people [for doing this], only technology can do that.[1]

Furthermore, according to the same expert, the national government does not focus on uncovering petty corruption because it is too costly:

> They are not looking for the small expenses or small corruption because it is expensive to develop this work. They prefer to focus on something that is going to represent billions of dollars because it is going to pay for the investment of putting a few people working on this.[2]

Anti-corruption efforts require resource-intensive tasks such as researching information, investigating suspicious patterns in datasets, or communicating corruption suspicions to the public (Adam & Fazekas, 2021). According to the interviewees, by taking over many of those tasks, AI can help governments and civil society organisations to allocate human and financial resources for anti-corruption efforts more efficiently, contributing to better corruption prevention, detection and/or prosecution. The interviewees' statements, resulting from practical work, are in line with previous theoretical research considerations. Sanchez-Graells (2024) confirms that AI-ACTs can support public officials in finding or aggregating information previously inaccessible or too costly to gather. In particular, AI-ACTs enable quick and continuous cross-checking of information. Without the use of AI-ACTs, such a profound analysis of large data sets would be nearly impossible to carry out manually (Köbis et al., 2022).

Along similar lines, another civil society actor added that AI could help to focus human resources. Since AI analyses all data, auditors can use their time more effectively and conduct targeted investigations:

> The [country] government is gigantic, it's impossible to audit everything. The main opportunity I see is to focus the eyes of the auditors, like "this is what you should focus on." If you look into everything, you look into nothing at the same time, because you cannot properly investigate everything that's on your table. That's the main opportunity I see.[3]

The interviewees agreed on the time- and cost-saving character of AI-ACTs. However, realising these positive contributions of AI can be challenging as AI-ACTs involve highly complex systems. Literature emphasises that not only the development of these tools, but also the application of them requires special expertise (Aarvik, 2019). A useful distinction draws on three categories of AI-driven jobs that are necessary for a successful implementation of AI-based technologies, namely (1) Human-AI Trainers, (2) Explainers, and (3) Sustainers (Wilson et al., 2017). In the first category, humans train AI-ACTs on how they should perform. They do this by programming algorithms on which the AI-based system should classify cases. The second category aims to close knowledge gaps and to provide clarity among non-technicians (Wilson et al., 2017). Explainers need to provide information on how the respective AI-ACTs arrive at a distinct classification. For instance, if an AI-based ACT flags a public procurement as risky, the explainer needs to be able to identify the indicators that led to this decision (for example, single bidding, no call for tender published, or a too-short advertisement period) (Fazekas & Kocsis, 2020). In the third category, humans as sustainers ensure that AI-based technologies operate as planned (Wilson et al., 2017). Sustainers address ethical as well as fairness issues. For instance, they are responsible for detecting and mitigating unintended obstacles, such as algorithmic bias.

Expert3 argues that AI offers innovative data analysis tools capable of reliably analysing large volumes of data to detect suspicious patterns. They compare the use of AI with traditional investigative journalism. Still, they point out a critical difference:

> The difference: We can analyse things on a much bigger scale than the usual investigative journalism that is looking into one case only. This is also very important, but with big data, we can look at the large patterns, over time and over a lot of countries, institutions and so on.[4]

All the above statements illustrate that AI-ACTs can facilitate more efficient use of scarce resources. At the same time, developing, implementing and maintaining such AI projects is expensive as they require investments of hardware, software, data storage and Human-AI experts. Thus, while AI-ACTs have great potential to improve the fight against corruption, such projects need to be carefully analysed, planned, and executed.

Increasing transparency

The interviews revealed that AI-ACTs have great potential to increase transparency. Expert1 ranked transparency as a top priority at every stage of the project:

> We wanted to start showing the population what the politicians were doing. It's more about accountability rather than judging.[5]

They further highlighted that commitment to transparency drove the team's decision about the technological design of the AI system. Instead of using complex neural networks, they opted for simpler machine learning models:

> We didn't want to create any "Black-Box". If we decided from the beginning to use artificial neural networks, we would probably be able to create something more powerful in a machine learning definition. It would be capable to get more cases to get less false positives. In some ways, it would be a better tool. But in the end, we wouldn't be able to explain what made a specific case suspicious. Of course, it would have been possible to do, but it would take some work to make this work transparent. So, we didn't want to focus on this technique.[6]

To further foster transparency, some projects publicised all of their data and code on platforms such as GitHub. They also provided technical documentation in English to encourage the government and civil society organisations to further develop the project or to undertake similar projects.

> We publish everything we do. So, we publish our codes on GitHub and we try to make everything replicable. So, you can download the raw datasets and you can download the codes that we used. Also, for example for the open tender portals, you can download the whole code that is behind the website. I think in that sense, we try to be accountable. Just as researchers try to be transparent.[7]

By publishing all codes publicly on GitHub, it becomes easier for potential collaborators to participate, to replicate the project in their own country, or to adapt the codes to their country's unique legal, political, and social circumstances (Savaget et al., 2019). Along similar lines, Expert4 added:

> The whole documentation of the [project name] tool is public, so there is not a single line of code that is not accessible meaning that if we would have working indicators or indicators that are working with algorithms, machine learning and so on, that would be public as well. Everyone could check themselves, what is processed and

> why a decision by the tool is made in that way. This is, what I think, one of the safeguards that needs to be ensured if you do anything based on AI in this sector.[8]

However, the high commitment to transparency also fulfilled the deliberate strategy of safeguarding the project against legal backlash:

> The whole project was about transparency. We wanted to be 100 per cent transparent rather than start being 10 per cent transparent and then work on improving. This was also a way of protecting ourselves because if we were sued by a politician, we would be able to go in front of a judge at any time and explain what happened.[9]

Using AI-ACTs helps to create transparency on multiple levels. First and foremost, they can detect corruption cases or suspicions that would otherwise remain hidden. Disclosing such cases to investigators and/or the public can result in reputational losses and trigger investigations or prosecutions, thereby increasing transparency. Furthermore, successful flagship AI-based projects might push other governments to publish more open data, which can then be used for anti-corruption efforts. Overall, AI-based projects show a big commitment to transparency by publicly making all data and code available. This openness is essential for building trust in such projects since they are by definition highly complex and challenging to grasp for lay people. The transparency commitment of those projects also has great potential to create ripple effects as they inform and hopefully motivate other projects and thereby advance technology-based anti-corruption efforts.

Enhancing civic engagement

The experts emphasised the importance of actively involving citizens in the political process. Expert1 highlighted that they strive to enable citizens to form their own opinion about the government's actions. Aspiring to an AI design that is understandable for lay people further reflects this pursuit of citizen involvement. Motivating citizens to participate in politics was the project's top priority. The project does not merely provide information to citizens, but invites them to act on it, such as further investigating by contacting members of parliament directly. Expert1 views citizens not as passive receivers but as active participants. Other experts similarly mentioned involving citizens in the political process as a key objective in their AI-ACT projects.

> For me, civic engagement, this is the most important part. Data and transparency without users are not worth anything. And this is why we are doing this. Because actually the intention of [project name] was to show people what their governments are doing, so you can go and check what happened in my city.[10]

> The idea came up that it would be an interesting tool to do some prefiltering and scanning of procurements to help ourselves and help journalists and other civil

> watchdogs to monitor procurement by pre-selecting those that are worth looking at.[11]

While some experts believe in the mobilisation potential of AI-ACTs, the way citizens currently make use of them does not live up to that potential. Expert3 explained that AI-ACTs efforts fail to mobilise citizens because they focus too much on data analysis and too little on usability.

> This is actually the crucial part that we sometimes forget because we are so stuck in data analysis. But this is what it's for in the end. It's quite difficult to see who is looking at this [e-procurement system]. Mostly it's only a few NGOs and the greater public is barely aware of public procurement and what is happening there, what kinds of corruption there are.[12]

They went on to say that harnessing the full mobilisation potential requires more media exposure for such tools. Furthermore, citizens' education about new technologies is lagging.

Expert5 emphasised the importance of the 'golden triangle of partnership' between the government, businesses, and the civil society for successfully monitoring public procurement. The project's platform calls upon all stakeholders in public procurement (citizens, companies, control and law enforcement authorities) to give feedback on public tenders. Expert5, therefore, repeatedly referred to the emergence of a 'monitoring community' as a result of the project.

From the statements of the experts, it emerges that AI-ACTs further contribute to mobilising citizens for anti-corruption efforts by balancing information asymmetry. That is, AI can provide information about corruption cases and motivate citizens to sign a petition, contact accused officials, or publicly speak out against corruption. Such tools, thus, give new ways to hold officeholders accountable and play a relevant watchdog role. However, previous research found digital anti-corruption tools are often used by very few people, particularly young and tech-savvy citizens (Martinez & Kukutschka, 2016). Often, citizens are neither aware that such tools exist, nor do they know how to use them. Thus, the development of AI-ACTs that are built to engage citizens should be accompanied by communication efforts such as marketing and public relations, aimed at raising public awareness, political efficacy, and ultimately public participation. Establishing civil society organisations in the anti-corruption community can play a vital role in this process. Moreover, developers need to take usability and user experience into account when designing such AI technologies. Thus, collaborations between computer scientists, social scientists and practitioners are vital for AI-ACTs to come to fruition.

Key Limitations

Insufficient quality and quantity of data

All interviewees agreed that accessing and obtaining relevant data presents a big problem for developing efficient AI-ACTs. Expert1 argued that, even though open government data initiatives and access to information laws publicise large datasets, they are usually not curated.

> I am a software engineer for some time and even for myself and for other people in the team, it was pretty hard to read the data from this file […] it was a file of 6GB. Your computer must have much more memory than the file to read it. You couldn't even open the file without having 16GB or more of that of RAM. So, you need a very expensive computer to even open the file.[13]

This quote suggests data availability cannot be equated to data accessibility as the latter requires sufficient hardware and software, and expert knowledge of how they work. However, Expert1 added that in their case the government swiftly responded to complaints and improved the datasets in terms of readability.

> Nowadays, and I could say that we were responsible for this change, the data is accessible. Because we met the people writing the software for releasing this file in the congress website. We showed them the needs of data scientists, of journalists like the format that they usually open, how they are going to open this file, what they are going to do with this file. Nowadays you don't even have to download the file for these expenses anymore.[14]

Other experts also addressed the issue of data accessibility. Expert3 and Expert4 both mentioned that the lack of sufficient data hindered their AI-ACT systems to realise their full potential. Another challenge consists of diverging standards in publicising data across countries. Some countries publish data in the CSV or JSON format, while others only publish data in PDF format. The latter required additional technical efforts by engineers.

> We quite soon understood that the data publication in TED [Tenders Electronic Daily] is also not perfect, so let's say if you take Germany, it's quite a mess how procurements are uploaded there and maintained. So, in this sense, Hungarian procurement in the TED are relatively well published I would say, surprisingly well published.[15]

Expert3 argued that the issue with data available in Germany is based on its federal structure.

> The difficulty about Germany is that it has a federal structure […] it depends on which Bundesland is doing what. There are definitely other countries in the EU

> that publish more and better than Germany. For example, in Slovakia, you have to publish everything, even if it's a contract of five euros. Portugal as well, I think. Also, in Germany, you don't have downloadable datasets as I remember. There is a lot more that needs to be done in Germany.[16]

Furthermore, two experts stressed that data from commercial registries bear great potential for AI-ACTs, but are often inaccessible.

> We were hoping that we would be able to connect the data with company registry data, but we were unable to. In Hungary, company registry data is not public as a database. You can individually look up companies for free, but you cannot scrape or not easily scrape the database, the companies consider it as stealing […] So, this is why we needed to give up on that very important element of the approach to have company data involved.[17]

Expert3 also emphasised that only a few countries publish commercial registry data and even if they do, integration of such data into internal data standards presents a challenge. However, they highlighted using such data to analyse ownership structures and filter out possible conflicts of interest as a fruitful addition. Expert2 also believes that adding additional data would lead to more accurate results by AI systems.

> Yes, there is. A person's assets. Do they have houses, apartments, boats etc. We would like to have information on that but that [data] is protected by fiscal laws and everything. Even though we work for the [country] government, we cannot just go there and grab the data we want. There are very strict privacy laws when it comes to that.[18]

Expert3 and Expert5 addressed another problem associated with data availability. Missing or incomplete data also represents a risk as data gaps can be used by officials to cover up corrupt transactions.

> Another disadvantage is that because we don't have some information in machine readable format, so some risks you can only see if you read documentations manually. If the procurement entity knows the risk indicators, he can try to make this tender look like good but hide all the information in the [manual] tender documentation. And we still don't have all fields that we need to analyse in electronic format. Therefore, someone can hide the information in the tender documentation and these indicators will show that the procedure is okay.[19]

According to the interviewees, the success of AI-ACTs hinges on data availability and data quality. Hence, to unleash their full potential, more access to and curation of data is needed. The interviews suggest that countries differ widely in terms of data availability and quality. This is consistent with the Global Open Data Index that is published by the Open Knowledge Foundation

(2016). Those countries, with high levels of corruption and high levels of open data, will benefit the most from AI-ACTs (Petheram, 2019). According to Petheram, those countries include Argentina, Brazil, Bulgaria, Colombia, Mexico, Paraguay, Romania, Slovakia, Russia, and Ukraine. Interestingly, all countries, mentioned in the report, are located in South America and Eastern Europe. These were also the two geographic areas in which our interviewees' AI-ACT projects were developed and executed. This shows that open data initiatives provide fertile ground for bottom-up AI-ACTs to emerge. This finding could encourage other countries to publish useful data instead of merely publishing unnecessarily large files of unreadable data. Large amounts of high-quality data are needed for AI-ACTs to make accurate decisions (Sanchez-Graells, 2024). In contrast, biased input data may produce wrongful accusations or simply useless results, true to the motto 'garbage in, garbage out'. This problem is particularly pronounced when AI agents act autonomously without much human oversight. Since suspicions or accusations of corruption come with a strong negative stigma for the accused, it is imperative for AI-ACTs to produce accurate decisions. Distributed ledger technologies, such as blockchain, have great potential to provide an immutable, transparent, and privacy-preserving data infrastructure (Aggarwal & Floridi, 2018). Thus, DLTs can complement existing and future AI-ACTs.

Legal restrictions

All respondents agreed that the legal framework is essential for the successful use of AI. They criticised that existing regulations are an obstacle for the most efficient rollout of their project. When asked what regulatory aspects would need to change for their project to have a greater impact, Expert2 replied that entirely new legal conditions would have to be created.

> I think that everything would have to change. For instance, instead of measuring the risk of corruption, after you enter the government, why not do that before you enter the government? Some features capture aspects of your life, that proceed your entrance in the government, such as how much money you've made so far, and what types of jobs did you have so far. And why not use that information before you employ someone into a public position? But right now, that's impossible, because that's prejudging the person.[20]

Expert3 also emphasised that their AI-ACT project would benefit from receiving data at an earlier stage.

> I wish we could do these analyses before or simultaneously. And I think that would be ideal. If there would be some automatic checks on the website, before a contract is actually signed. But usually we work after the fact, because this is when the data gets published. You might see a call for tender and you might see who's bidding for

> it obviously, but usually we get the information later because that's when governments upload it.[21]

Along these lines, Expert5 mentioned that their team is trying to work even more closely with the public authorities and ministries to ensure that public institutions follow up on the suspicions the project reveals.

> We have the same problems, that there are no implications of our project, due to existing law. That's why we try to work more closely with our controlling authorities and our ministry to help that if we find something, then we can ensure that this finding will make some changes or that someone will be punished or their controlling authority will take actions and go to court and the procurement authority will be punished for what it did. Or if there is a collusion between the bidders, then the bidders will get punished and therefore won't violate the law in the future.[22]

Expert1 expressed the need for changes in the legal framework to make AI systems more effective against corruption. They mentioned three necessary steps: first, governments need to be obliged to actively publish the relevant data. This right must have constitutional status in the respective country. Second, it should be possible to process all data through algorithms without having to fear legal ramifications. Here, they referred to the risk of being exposed to justiciable accusations of defamation by politicians. Third, the infrastructure for sufficient funding for AI-ACT projects needs to be extended. According to Expert1, if reliable government bodies pursued AI-ACT projects, the problem of insufficient funding could be somewhat alleviated. They emphasised that the decision to start a crowdfunding campaign was the only way to get their project off the ground. However, Expert2 rejects the third assertion, arguing that governments could abuse their power if entrusted with AI-ACTs.

> In terms of risk, I think the government is the main risk. The risk is, a government steps in and decides to control things like you know, I'm the regulator, I'm going to tell you, what you can do, what you can automate, what sorts of models you can build, and which ones, you cannot build. Because it's very tempting for the government.[23]

Similarly, Expert3 had some ideas for improving the legal framework. They emphasised the importance of the data availability and argued that governments should be required to proactively publish data. In particular, they noted that such open government data must be disclosed in a structured and machine-readable form.

> It would be great if countries would have automated checks themselves on procurement. So, they could use this kind of method that we use to check themselves and

> to monitor continuously and to have it built in the system instead of us taking the information, analysing it and telling them what looks fishy. It would be great if it would be built into their system and if the authorities in the countries would actually use these methods. But, in most countries, it's not even mandatory to publish everything. So, we are a long way away from that.[24]

The interviewees emphasised that unclear regulatory frameworks do affect not only the access to data, but also the subsequent implications of corruption allegations exposed by AI-ACTs. The experts in our sample criticised that law enforcement agencies often fail to act on corruption suspicions. Thus, in many cases, AI-ACT efforts do not have tangible consequences for corrupt officials. Therefore, more collaborations between AI-ACT projects and public institutions are needed. For instance, public administrations could organise hackathons and issue seed-funding grants for promising bottom-up AI-ACTs. Such an approach could help create a collaborative environment between civil society actors, public officials, and researchers.

Technological challenges

Besides the immense potential of AI for anti-corruption, the experts also acknowledged the technological challenges associated with using machine learning to identify corruption. For instance, Expert2 mentioned the risk of wrongful accusations due to biased input data.

> Of course, any model has its biases. Usually, the public servants that have been expelled before, they're low rank. It's not the fat cats on the top of the pyramid. We don't have any ministers there. That introduces a bias in the model because the model hasn't seen a lot of ministers that have been expelled before. So, the model kind of learns incorrectly that if the salary is low, the probability of corruption is higher. That's a problem for us.[25]

Creating unbiased databases is challenging, especially in the context of anti-corruption efforts. For instance, it often takes a very long time from an accusation of corruption to a legally binding conviction. This may lead to the problem that a corruption case, that is only classified as corrupt at a later point in time, is still marked as non-corrupt in the data set (Sanchez-Graells, 2024). Moreover, corruption is difficult to assess and to quantify by its very nature as it happens in secret. In a high corruption context, data might further be biased due to manipulation by corrupt actors, tinkering with the database to hide suspicious activities. Expert2 further stresses that AI is never used to make final decisions without human oversight.

> We don't worry too much about that because there is an investigation afterwards and if there is no corruption, then the person will be exonerated. Nothing bad will happen to them. It's not like people get fired because of the model. It's not

> something we are too concerned about right now, maybe in the future, if the model becomes automatically enforceable. Like if you have a probability of corruption, higher than point 7, then you are fired. Then, this will become a problem and it will become something we have to worry about. But right now, if you have nothing to hide, you have nothing to fear.[26]

Another challenge addressed by Expert4 and Expert5 refers to training algorithms with labelled data, hence using supervised machine learning. Both mentioned that creating and updating training databases is a challenge.

> We did some machine learning but rather in an experimental way. Because the difficulty we were facing in terms of machine learning is building test databases and teaching databases in order to clearly identify corrupt procurements […] the teaching database that we have done could not become as big that we could rely on the patterns it produced. So, we stopped this approach some time ago.[27]

> The problem is that we need a lot of comparisons for us to teach the system. But we don't have so many experts that can decide which tender is more risky, based on the parameters. So, we don't have enough training examples to make it run smoothly and we don't know how to get enough training examples […] It's multiple problems. In the future, [project] will add new and new parameters and every time we need to reteach this system and we always need these people to provide their time to teach the system. The idea behind ML is good but we need a lot of human resources and training examples. That's why we are currently on pause with this ML.[28]

Expert3 also touched on another aspect related to the complexity of AI-ACTs, namely that such projects are too technical and difficult to apply:

> What we hear a lot is that it is too technical and that it is hard for people to understand and to use […] I think this is one huge challenge. Maybe we need more media and more civil engagement to make it understandable for people. I think that is one big criticism. We are academics and researchers and I think that is where we usually fail, to convey a message to the general public.[29]

Opacity, resulting from the complexity of AI-ACTs, is a well-acknowledged risk among scholars (Ponti et al., 2021). The statement of Expert3 indicates that interdisciplinarity within AI-ACT projects seems of great importance.

CONCLUSION

Based on five semi-structured interviews with experts involved in projects using AI in anti-corruption efforts, this study explores the benefits and risks of such AI-based technologies. The results reveal that civil society actors see both potential and limitations of AI in the fight against corruption. Some aspects prominently mentioned in the interviews also apply to designing and implementing AI in other societal areas. This includes the hope to allocate human and financial resources efficiently. Also, the limitations of an unclear and often insufficient regulatory framework do not exclusively relate to AI-ACTs but other AI systems as well.

However, the interviews also carved out some corruption-specific insights. For instance, anti-corruption efforts require an active civil society holding political elites accountable.

A fruitful approach might be to combine expertise from different fields, such as computer science, communication science or psychology, to develop more effective methods to convey complex knowledge to the public. To tackle opacity, it is essential to educate the public about which algorithms are used, how they function and which data they are based on. Such transparency could be a major step towards establishing trust. A successful example for an interdisciplinary AI-ACT is the Brazilian project *Operation Serenata de Amor*.[30] It uses an automatic algorithm to analyse financial expenditures of Brazilian members of parliament. The project was founded by a team with expertise in computer science, business and sociology. In addition, government and civil society organisations could jointly offer courses in which citizens are taught programming languages (Ponti et al., 2021). By reducing opacity and at the same time providing practical knowledge, the barriers to civic participation would be lowered.

Consequently, bottom-up AI-ACTs often face an intricate two-step challenge: first, the need to disclose corruption and second, the need to engage citizens to act on the disclosed information. This makes AI-ACTs arguably particularly complex to design and implement. Moreover, as corrupt acts are per definition hidden from plain sight, aggregating useful data and establishing a ground truth is likely to be more difficult with regard to corruption.

To conclude, we hope that the insights gained in this chapter can inform developers and anti-corruption practitioners, to team up to leverage AI in the fight against corruption.

NOTES

1. Interview with Expert1 conducted on 06.09.19, online.
2. Interview with Expert1 conducted on 19.09.19, online.
3. Interview with Expert2 conducted on 06.09.19, online.
4. Interview with Expert3 conducted on 17.07.20, online.
5. Interview with Expert1 conducted on 19.09.19, online.
6. Interview with Expert1 conducted on 19.09.19, online.
7. Interview with Expert3 conducted on 17.07.20, online.
8. Interview with Expert4 conducted on 21.07.20, online.
9. Interview with Expert1 conducted on 19.09.19, online.
10. Interview with Expert3 conducted on 17.07.20, online.
11. Interview with Expert4 conducted on 21.07.20, online.
12. Interview with Expert3 conducted on 17.07.20, online.

13. Interview with Expert1 conducted on 19.09.19, online.
14. Interview with Expert1 conducted on 19.09.19, online.
15. Interview with Expert4 conducted on 21.07.20, online.
16. Interview with Expert3 conducted on 17.07.20, online.
17. Interview with Expert4 conducted on 21.07.20, online.
18. Interview with Expert2 conducted on 06.09.19, online.
19. Interview with Expert5 conducted on 27.11.19, online.
20. Interview with Expert2 conducted on 06.09.19, online.
21. Interview with Expert3 conducted on 17.07.20, online.
22. Interview with Expert5 conducted on 27.11.19, online.
23. Interview with Expert2 conducted on 06.09.19, online.
24. Interview with Expert3 conducted on 17.07.20, online.
25. Interview with Expert2 conducted on 06.09.19, online.
26. Interview with Expert2 conducted on 06.09.19, online.
27. Interview with Expert4 conducted on 21.07.20, online.
28. Interview with Expert5 conducted on 27.11.19, online.
29. Interview with Expert3 conducted on 17.07.20, online.
30. See Chapter 2 in this volume for further applications of digital technologies, including AI, in Brazilian anti-corruption and pro-accountability civil society initiatives.

REFERENCES

Aarvik, P. (2019). Artificial Intelligence a promising anticorruption tool in development settings. CMIU4 AntiCorruption Resource Centre. https:// www .u4 .no/ publications/artificial-intelligence-a-promising-anti-corruption-tool-in-development -settings.pdf.

Adam, I., & Fazekas, M. (2021). Are emerging technologies helping win the fight against corruption? A review of the state of evidence. *Information Economics and Policy, 57,* 1–14. https://doi.org/10.1016/j.infoecopol.2021.100950.

Aggarwal, N., & Floridi, L. (2018). The opportunities and challenges of blockchain in the fight against government corruption. 19th General Activity Report, Council of Europe Group of States Against Corruption (GRECO).

Ala-Pietilä, P., Bonnet, Y., Bergmann, U., Bielikova, M., Bonefeld-Dahl, C., Boujemaa, N., Bauer, W., Bouarfa, L., Chatila, R., Coeckelbergh, M., Dignum, V., Gagné, J.-F., Goodey, J., Haddadin, S., Hasselbalch, G., Heintz, F., Hidvegi, F., Höckner, K., Jégo-Laveissière, M.-N., … Van Wynsberghe, A. (2019). *Building Trust in Human-centric AI*. AI HLEG.

Ash, E., Galletta, S., & Giommoni, T. (2020). A machine learning approach to analyzing corruption in local public finances. *Center for Law & Economics Working Paper Series*, *6*.

Choi, D., & Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. *Security and Communication Networks*, 4, 1-15. https://doi org/10.1155/2018/5483472.

Colliri, T., & Zhao, L. (2019). Analyzing the bills-voting dynamics and predicting corruption-convictions among Brazilian congressmen through temporal networks. *Scientific Reports* 10.1038/s41598-019-53252-9.

de Blasio, G., D'Ignazio, A., & Letta, M. (2020). Predicting corruption crimes with machine learning. A study for the Italian municipalities *Working Paper Series, 16/2020.*

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78- 87. https://doi.org/10.1145/2347736.2347755.

European Commission. (2021). *ARACHNE risk scoring tool*. https:// ec .europa .eu/ social/main.jsp?catId=325&intPageId=3587&langId=en.

Fazekas, M., & Kocsis, G. (2020). Uncovering high-level corruption: Cross-national objective corruption risk indicators using public procurement data. *British Journal of Political Science*, *50*(1), 155–164. https://doi.org/10.1017/S0007123417000461.

Global Witness (2021). How can we use artificial intelligence to help us fight corruption in the mining sector? https://www.globalwitness.org/en/blog/how-can-we-use-artificial-intelligence-help-us-fight-corruption-mining-sector/.

Köbis, N., Starke, C., & Rahwan, I. (2022). The promise and perils of using artificial intelligence to fight corruption. *Nature Machine Intelligence, 4*, 418–424. https://doi .org/10.1038/s42256-022-00489-1.

Kossow, N., & Dykes, V. (2018). Embracing digitalisation: How to use ICT to strengthen anti-corruption. *GI Anti-Corruption and Integrity Programme*. https:// www.giz.de/de/downloads/giz2018-eng_ICT-to-strengthen-Anti-Corruption.pdf.

Li, X., & Juhola, M. (2015). Country crime analysis using the self-organising map, with special regard to economic factors. *International Journal of Data Mining, Modelling and Management*, *7*(2), 130–153. https://doi.org/10.1504/IJDMMM.2015.069250.

Lima, M. S. M., & Delen, D. (2020). Predicting and explaining corruption across countries: A machine learning approach. *Government Information Quarterly*, *37*(1), 101407. https://doi.org/10.1016/j.giq.2019.101407.

Littig, B. (2009). Interviewing the elite – interviewing experts: Is there a difference? In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts* (pp. 98–113). Palgrave Macmillan UK. https://doi.org/10.1057/9780230244276_5.

López-Iturriaga, F. J., & Sanz, I. P. (2018). Predicting public corruption with neural networks: An analysis of spanish provinces. *Social Indicators Research*, 140(3), 975-998. https://doi.org/10.1007/s11205-017-1802-2.

Martinez, R., & Kukutschka, B. (2016). Technology against corruption: the potential of online corruption-reporting apps and other platforms, U4 Expert Answer, 20. https:// www .u4 .no/ publications/ technology -against -corruption -the -potential -of -online -corruption-reporting-apps-and-other-platforms.

Mattoni, A. (2021). Digital media in grassroots anti-corruption mobilizations. In D. A. Rohlinger & S. Sobieraj (Eds.), *Sociology and Digital Media*. Oxford University Press. 10.1093/oxfordhb/9780197510636.013.46.

Mayring, P. (2015). Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping, & N, Presmeg (Eds.) *Advances in Mathematics Education* (pp. 365–380). https:// doi .org/ 10 .1007/ 978 -94 -017 -9181 -6_13.

Mazrekaj, D., Titl, V., & Schiltz, F. (2023). Identifying Politically Connected Firms: a machine learning approach. *Oxford Bulletin of Economics and Statistics*, 86(1), 137-155. https://doi.org/10.1111/obes.12586.

Mhlanga, D. (2021). Financial inclusion in emerging economies: The application of machine learning and artificial intelligence in credit risk assessment. *International Journal of Financial Studies*, *9*(3), 39. https://doi.org/10.3390/ijfs9030039.

Microsoft (2021). Microsoft collaborates with the Nigerian government to accelerate digital transformation in the country . .https:// news .microsoft .com/ en -xm/ 2021/

05/ 03/ microsoft -collaborates -with -the -nigerian -government -to -accelerate -digital -transformation-in-the-country/.
Miles, J., & Gilbert, P. (2005). *A Handbook of Research Methods for Clinical and Health Psychology*. Oxford University Press. https:// doi .org/ 10 .1093/ med: psych/ 9780198527565.001.0001.
Mungiu-Pippidi, A., & Heywood, P. (2020). *A Research Agenda for Studies of Corruption*. Edward Elgar Publishing. https://doi.org/10.4337/9781789905007.
Odilla, F. (2023). Bots against corruption: Exploring the benefits and limitations of AI-based anti-corruption technology. *Crime, Law and Social Change*, 80(4), 353-396. https://doi.org/10.1007/s10611-023-10091-0.
Olszewski, D. (2014). Fraud detection using self-organizing map visualizing the user profiles. *Knowledge-Based Systems, 70,* 324–334. https://doi.org/10.1016/j.knosys .2014.07.008.
Open Knowledge Foundation. (2016). *Global Data Index*. https:// index .okfn .org/ place/
Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, (5), 533–544. https://doi.org/10.1007/s10488-013-0528-y.
Petheram, A. (2018, July 27). From Open Data to Artificial Intelligence: The Next Frontier in Anti-Corruption. *Oxford Insights*. https:// www .oxfordinsights .com/ insights/aiforanticorruption.
Petheram, A., Pasquarelli, W., & Stirling, R. (2019). The next generation of anti-corruption tools: big data, open data & artificial intelligence. *Oxford Insights*. https://www.oxfordinsights.com/ai-for-anti-corruption.
Ponti, B., Cerrillo-i-Martínez, A., & Di Mascio, F. (2021). Transparency, digitalization and corruption. In E. Carloni & M. Gnaldi (Eds.), *Understanding and Fighting Corruption in Europe.* Springer. https://doi.org/10.1007/978-3-030-82495-2_6.
Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A., … Wellman, M. (2019). Machine behaviour. *Nature*. https://doi.org/10.1038/s41586-019-1138-y.
Rothstein, B. (2011). *The Quality of Government: Corruption, Social Trust, and Inequality in International Perspective*. University of Chicago Press.
Sanchez-Graells, A. (2024). Procurement corruption and artificial intelligence: Between the potential of enabling data architectures and the constraints of due process requirements. In S. Williams-Elegbe & J. Tillipman (Eds.), *Routledge Handbook of Public Procurement Corruption*. Routledge, Forthcoming.
Savaget, P., Chiarini, T., & Evans, S. (2019). Empowering political participation through artificial intelligence. *Science and Public Policy*, *46*(3), 369–380. https:// doi.org/10.1093/scipol/scy064.
Swiderski, B., Kurek, J., & Osowski, S. (2012). Multistage classification by using logistic regression and neural networks for assessment of financial condition of company.
Decision Support Systems, *52*(2), 539–547. https:// doi .org/ 10 .1016/ j .dss .2011 .10 .018.
Wilson, H. J., Daugherty, P., & Morini-Bianzino, N. (2017). The jobs that artificial intelligence will create. *MIT Sloan Management Review*, *58*(4), 14–16.

World Bank. (2020). *Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks*. World Bank Publishing. http:// hdl .handle .net/ 10986/35317.

Wroblewski, A., & Leitner, A. (2009). Between scientific standards and claims to efficiency: Expert interviews in programme evaluation. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts* (pp. 235–251). Palgrave Macmillan UK. https://doi.org/10.1057/9780230244276.