Open Data in Elections

Why It Is Needed and What Can Be Done to Make More Progress

Peter Wolf

Introduction

The discussion on how technology can be applied to improve the integrity and transparency of elections often focuses on complex and expensive technologies such as biometrics (Wolf et al. 2017), electronic voting (Wolf, Nackerdien and Tuccinardi 2011) and voter identification systems. Something unfortunately mentioned far less often is the potential to apply technology in a way that is not only much cheaper, but also arguably more efficient—the use of online tools that provide access to election-related information based on open data principles. This is indeed unfortunate as the absence of at least rudimentary open data approaches in elections can undo and even reverse the integrity gains brought about by other technologies and measures.

The case of the Democratic Republic of the Congo

A case in point were the 2018 elections in the Democratic Republic of the Congo (DRC). Starting with biometric voter registration in 2005–2006 (International IDEA and RECEF 2018), election technology was successively introduced and expanded, culminating in the deployment in 2018 of hotly debated electronic voting machines. These voting machines were in fact treated as 'ballot printing machines' in order to fit within the country's legal framework (Congo Research Group 2018).

Given the extremely harsh environmental conditions, limited infrastructure and the otherwise limited technological penetration in large parts of the DRC, the deployment and use of these technologies by voters worked surprisingly well. Millions of voters cast their ballots facilitated by technology. Thousands of international (Carter Center 2018) and domestic observers (SYMOCEL 2019) monitored the process throughout the country.

However, it is arguable that what followed election day was where things went badly wrong. The results based on the voting machines reached the central election administration within a few hours or days, and vote totals were published and a winner announced a few days later. However, some observers expressed concern that the results they had observed in the polling stations and tallied in parallel did not necessarily match the published results. There is usually a way to resolve such claims. Observers can simply compare their observations and data from the polling stations with the results published by the election administration. In this case, however, detailed polling station results were never published and observers were not given an opportunity to check for any discrepancies.

Fortunately, the large-scale election violence that was feared by many did not take place. While doubts about the election result remained, a political solution was eventually found—but a stain on the integrity of the election

remains. After many years and a double- to triple-digit million dollar spend on massive technology upgrades for every election, the bottom line was an election that not only missed its goal of increasing transparency but actually achieved the opposite—a massive information disbalance. The election results were available faster than ever before but only for a small and select group that was able to gain early insights and prepare accordingly. At the same time, the general public and most electoral stakeholders have been kept in the dark to this day.

Overview

To avoid any similar disappointing outcomes in the future, open election data should be made a central element of electoral transparency initiatives and any technology upgrades for elections. In fact, very little technology is needed for open data. Even a mostly paper-based election can be made much more transparent with the addition of a few relatively simple open data systems.

Building on the International IDEA publication *Open Data in Electoral Administration* (Carolan and Wolf 2017), this paper provides an overview of what open data is, where it can be applied in elections, how progress can be made, and the roles and responsibilities of electoral management bodies (EMBs) and other stakeholders. It also addresses a set of risks that must be addressed in any open election data project.

What is open election data?

The concept of open data is based on the idea that data should be freely available for anyone to access, use and share. Providing election data as open data has numerous benefits. It provides civil society, citizen journalists, electoral observers and citizens with access to the same detailed information that was previously only available to selected stakeholders such as large media outlets. In doing so, open data allows all interested stakeholders to follow and understand the electoral process and can lead to more inclusive, transparent and trusted elections. The data can be used in many different ways, such as to feed news websites and develop voter information applications, and for in-depth data analysis and investigations.

The Open Election Data Initiative (Open Election Data Initiative n.d.) establishes nine principles that can guide the release of open election data and support this deliberation process. According to the nine principles, open election data should be:

1. timely: made available as quickly as necessary to be useful;
2. granular: available at the finest-possible level or granularity or detail and also made available at the primary level, meaning the level at which the source data was collected;
3. available free of charge on the Internet: released without any monetary restrictions;
4. complete and in bulk: released as a complete data set without any omissions;
5. analysable: made available in a machine-readable format that can be quickly and easily analysed;
6. non-proprietary: available in a format over which no entity has exclusive control;
7. non-discriminatory: available to any individual or institution for anonymous access without any usage restrictions, including application or registration requirements;
8. licence-free: there should be no barriers to reuse or redistribution for any purpose;
9. permanently available: at a stable Internet location for an indefinite period of time.

Specifically, data that is only available for a short period of time is not open.

Where can open election data principles be applied?

As noted in the DRC example above, election results are an obvious data set that can be published according to open data principles. However, there are more opportunities and open data concepts can be applied in all phases of the electoral cycle:

1. Boundary delimitation: granting access to population data, as well as the spatial data for electoral districts derived from the population data (data tables and maps).
2. Political party and candidate registration: granting access to lists of parties and candidates that applied, were registered and were rejected by the EMB, also including registration information.
3. Campaign finance: information about the funds parties and candidates received for and spent on their campaigns, including limits and regulations.
4. Voter registration: statistics about registered voters, including breakdowns by gender, age and geography, also including information about the registration process itself.
5. Voter lists: access to detailed information about eligible voters, including names, dates of birth, national identity numbers, residence, polling station and so on.
6. Polling stations: including addresses, contact information and number of voters for each polling station.
7. Election results: including the number of registrants, and valid and invalid votes, as well as the votes for each party and candidate. All this data should be available disaggregated down to the lowest level at which votes are counted.
8. Legal framework: granting access to all laws, regulations and instructions related to the electoral process (textual data).
9. EMB: access to contact information and details of officials at all levels (textual data).
10. EMB processes: data on EMB decisions, minutes and resolutions.
11. Election campaigns: data on the timetables for campaigns and related regulations and restrictions, as well as public resources for campaigns.
12. Voter education: access to the information made available for voters about political parties, candidates, and registration and voting procedures.
13. Electronic voting and counting: provision of details regarding the procurement, accuracy, security and reliability of the voting system, including source codes.
14. Electoral complaints and dispute resolution: information about the number and types of complaints, the entities filing them and the outcomes of dispute resolution, as well as details on accessing the complaints process.
15. Election security: information from regulatory, security and judicial institutions about standards for the police and military in the electoral process.

How can progress be made?

Government commitment

As of December 2020, 75 states, 14 of which are in the Asia-Pacific region, have made an in principle commitment to open data by joining the Open Government Partnership (Open Government Partnership n.d.) and endorsing the Open Government Declaration (2011). In recent years, the amount of election data openly published has been increasing around the world. In most countries, however, including many that have endorsed the Open Government Declaration, significant improvements are required with regard to the availability of open election data.

The role of civil society

Where open election data gaps exist, civil society can play an important role. This often happens in the form of initiatives that take whatever information is available from the election administration and convert it into more open formats. The other, arguably more important, role of civil society organizations (CSOs) is advocacy and encouraging the election administration to directly publish more open data. A successful advocacy approach has two benefits: the data adheres to the 'primary' principle as it comes directly from the original data source, the EMB; and CSOs are able increasingly to focus their resources on data analysis instead of on making data available.

The role of the electoral management body

Advocacy of open election data is important because some EMBs are still unaware of the potential usefulness and efficiency of open data for increasing electoral integrity. External motivation is also important as a move to open data is more than a technical update. It requires a cultural change for the institution that otherwise might not happen.

Generic guidance and recommendations exist on the technical process of introducing open data systems in elections, for example in Broad et al. (2015) published by the Open Data Institute. This guide highlights, among other things, the need to build support for and consensus on the initiative and to identify ways to overcome resistance to change.

In addition, disclosing election data requires the development of a culture of openness within the EMB. This culture entails a commitment to transparency combined with a recognition of the potential for data misuse and of the risks related to disclosing data.

Creating a culture of openness

Open data changes an EMB from a data owner into a data provider. As a data provider, an institution becomes more transparent and this opens up new avenues for feedback and criticism that need to be addressed accordingly. Mistakes happen in any institution and higher levels of transparency mean that they are exposed more easily and more quickly. When publishing large and detailed data sets, especially in real time and with necessarily limited quality checks, there is always a possibility that some mistakes will remain undetected before publication. A few exposed mistakes can create the impression that there are more, still undetected, problems in the electoral process.

Some actors can also be expected to look for such mistakes in an attempt to undermine or discredit the organization. If such situations are not handled well, the result can be the opposite of what was intended—instead of increasing integrity, mistrust in institutions and processes may be increased. This needs to be addressed in several ways. First, the EMB must proactively disclose and communicate any limitations of the open data they provide, as far as possible in advance to avoid misunderstandings that may arise. In addition, together with the release of data, ethical use of the data provided should be promoted. Where data is preliminary or partial, this must be clearly explained. Most importantly, the EMB must be highly responsive to complaints, requests for corrections and feedback from the public, political parties and others.

Managing the risks of open data in elections

The potential contribution of open data to electoral processes must therefore be balanced with an understanding of some of the sensitivities and risks related to the release of data.

Balancing transparency and privacy

In common with most sectors, some data pertaining to elections could be sensitive and might need to remain closed, to be shared with only a few parties or to be disclosed only above a certain level of aggregation. This is especially true where personal data is concerned, such as from voter registers, for which acceptable formats for publication and a balance between data protection and transparency will need to be carefully evaluated and managed.

The risk of misuse of data

In the light of a seemingly growing global trend towards attempts to undermine electoral integrity, it is important to address the risk that detailed election data could be misused, misinterpreted or misrepresented. As demonstrated in the recent US elections, a changed outcome between early polling station data and the final results can contribute to mistrust in the results process that persists beyond the period of uncertainty following the election. In the context of Covid-19, this period of uncertainty could get longer as many countries adopt or increase the use of special voting arrangements such as early, mobile and postal voting (Carolan and Wolf 2017). These generally lead to longer vote counting, results tabulation and announcement processes.

EMBs must therefore ensure that every effort is made to provide accurate information in order to counter misinterpretations. This includes the proactive provision of information on data release cycles and explanations of what to expect from and how to interpret early data.

These open data risks can be further mitigated through careful consideration of and planning for decisions to publish data, which should include consultations with those who wish to use the data and those who understand the sensitivities and risks involved.

Stakeholder responsibilities

Overall, however, it should be acknowledged that an open data initiative will not on its own provide a 'magic bullet' for improving the functioning of electoral processes, in particular making them automatically more transparent and trusted. An open data initiative is a vital factor in increasing transparency but is most impactful when combined with other integrity measures.

It is also important to note that open data initiatives, especially in their early stages, involve give and take between the election administration and electoral stakeholders, as well as responsibilities on both sides. Election administrators will need to embark on a sometimes difficult journey towards a culture of openness. Other electoral stakeholders, from political candidates and civil society entities to media outlets and citizens, will have to respond to this by ensuring that they fully understand and can analyse the data they receive, that they report and present it accurately—and that they call out and discipline rogue actors that misuse and misinterpret the information provided.

Conclusions

Among the various technical solutions for increasing electoral integrity and transparency, the application of open data principles and the implementation of related tools are arguably the most effective and one of the cheapest options available. At the same time, opening up election-related data requires not just a technical, but also a cultural change. Such change is neither simple nor risk free. Careful planning that takes global experience into account is therefore essential when embarking on open data projects in elections.

Much of this paper is based on International IDEA's 2017 guide Open Data in Electoral Administration (Carolan and Wolf 2017), which contains much more detail and provides references to further resources and expert organizations in the field of open data.

About the author

Peter Wolf is Senior Expert on ICT, Elections and Democracy at International IDEA.

Contact us

International IDEA
Strömsborg
SE–103 34 Stockholm
Sweden
Email: info@idea.int
Website: ‹https://www.idea.int›

References

Broad, E. et al., *Open Data in Government: How to Bring About Change* (London: Open Data Institute, 2015), <https://theodi.org/article/open-data-in-government-how-to-bring-about-change/>, accessed 18 March 2021

Carolan, L. and Wolf, P., *Open Data in Electoral Administration* (Stockholm: International IDEA, 2017), <https://doi.org/10.31752/idea.2017.5>

Carter Center, *Democratic Republic of the Congo 2018 Harmonized Presidential, Parliamentary and Provincial Elections*, Expert Mission Report, 2018, <https://www.cartercenter.org/resources/pdfs/news/peace_publications/election_reports/drc-2018-election-report-final-en.pdf>, accessed 18 March 2021

Congo Research Group, 'The electronic voting controversy in the Congo', Election Brief No. 1, April 2018, Centre on International Cooperation, New York University, <http://congoresearchgroup.org/wp-content/uploads/2018/04/Electronic-Voting-Controversy-1.pdf>, accessed 18 March 2021

International IDEA and Réseau des compétences électorales francophones (RECEF), *The Use of New Technologies in Electoral Processes, Workshop report: Praia, Cabo Verde, 22–23 November 2017* (Stockholm: International IDEA and RECEF, 2018), <https://www.idea.int/publications/catalogue/utilisation-des-nouvelles-technologies-dans-les-processus-electoraux?lang=en>, accessed 31 March 2021

Open Election Data Initiative, 'Section 2: Open election data principles', [n.d.], <https://openelectiondata.net/en/guide/principles>, accessed 18 March 2021

Open Government Partnership, 'Members', [n.d.], <https://www.opengovpartnership.org/our-members>, accessed 18 March 2021

—, 'Open Government Declaration', September 2011, <https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/>, accessed 18 March 2021

Synergie des Missions D'observation Citoyenne des Elections (SYMOCEL), *Rapport Final D'Observation des Elections Directes et Indirectes de 2018 et 2019* [Final Observation Report of the Direct and Indirect Elections of 2018 and 2019] (Kinshasa: SYMOCEL, May 2019), <https://www.eisa.org/pdf/drc2019symocel.pdf>, accessed 18 March 2021

Wolf, P., Nackerdien, R. and Tuccinardi, D., *Introducing Electronic Voting: Essential Considerations* (Stockholm: International IDEA, 2011), <https://www.idea.int/publications/catalogue/introducing-electronic-voting-essential-considerations>, accessed 18 March 2021

Wolf, P. et al., *Introducing Biometric Technology in Elections* (Stockholm: International IDEA, 2017), <https://www.idea.int/publications/catalogue/introducing-biometric-technology-elections>, accessed 18 March 2021