Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

December 2024

Table of contents

Forewords

I am pleased to present these Guidelines for AI in parliaments, which arrive at a crucial moment in our democratic journey. We stand at the threshold of a transformation that is reshaping how parliaments operate and serve their citizens. Artificial Intelligence presents both extraordinary opportunities and significant challenges for our institutions of democracy.

These Guidelines emerge from our recognition that parliaments must take a leading role in governing the use of AI, not only through legislation and oversight but also through their own adoption and implementation of these technologies. The Guidelines represent a collaborative effort, drawing on the expertise and experience of parliamentary staff and technology specialists from across our global community.

The IPU's commitment to supporting parliaments in their digital transformation journey has never been more vital. We continue to witness first-hand how technological advancement is reshaping parliamentary work. These Guidelines build upon our existing frameworks, including the IPU resolution on "The impact of artificial intelligence on democracy, human rights and the rule of law" and the IPU Charter on Ethics of Science and Technology (October 2024), as well as the Guide to Digital Transformation in Parliaments and the World e-Parliament Reports developed through our Centre for Innovation in Parliament.

The IPU resolution highlights both the tremendous potential of AI to enhance parliamentary functions and the critical importance of ensuring its responsible deployment, particularly in protecting our most vulnerable citizens. These Guidelines therefore emphasise the fundamental principles of transparency, accountability and fairness that must underpin any technological advancement in our institutions of democracy, as well as more technical aspects.

The Guidelines offer a comprehensive framework for parliaments at all stages of their AI journey, whether they are just beginning to explore AI's potential or are already implementing advanced applications. They provide practical guidance while emphasising the importance of strong governance, ethical considerations and risk management. Most importantly, they stress that AI should augment and enhance human capability rather than replace it, particularly in the context of democratic deliberation and decision-making.

These Guidelines are valuable for different audiences inside and outside of parliament, including for members, especially those serving on modernization committees, technology committees or committees of the future, as well as senior managers and technical experts. They offer useful insights for parliamentarians as they grapple with AI oversight and regulation in society, showing how AI can be deployed responsibly within their own institutions. They provide a framework for making informed decisions about AI adoption while ensuring robust democratic oversight.

Working together, we can ensure that AI serves to strengthen rather than diminish our democratic institutions, upholding the fundamental values that our parliaments represent.

Martin Chungong
Secretary General
Inter-Parliamentary Union

Since the Brazilian Chamber of Deputies' first experience with artificial intelligence (AI) in 2013, we have been on a journey of continuous learning about this technology and the extraordinary capabilities that it can offer to parliaments. Despite this, we are still surprised by the exponential speed of advances in AI, and by its ubiquitous nature in the daily lives of public organizations around the world.

Among the lessons learned over the years of using AI in the legislative branch, it is worth highlighting the need for coordination encompassing multiple stakeholders, so that the technology's use can be well planned and managed. This observation inspired us to hold the Parliamentary Data Science Hub meeting in Brasília (in April 2024) to discuss good practices for the use and development of artificial intelligence in parliaments.

I am proud to note that the excellent work of the experts who participated in that meeting has resulted in a set of guidelines with a user-friendly approach and the flexibility needed to meet the different realities faced by parliaments. The guidelines combine strategic actions and policies with examples of everyday practices, in a way that recognizes the plurality of decisions involved in the use, development and outsourcing of AI systems.

In addition to being an everyday reference document for IT professionals, the *Guidelines for AI in Parliaments* are a reference for parliamentary leaders, directors general, secretaries general and managers of parliaments, in defining strategies and priorities relating to this topic. I hope that these guidelines will be instrumental in reducing the technological gap between parliaments worldwide, so that they can keep up with the advances that society expects.

Celso de Barros Correia Neto
Director General
Chamber of Deputies of Brazil

Coordinator of the Parliamentary Data Science Hub in the IPU's Centre for Innovation in Parliament

Introduction

Artificial intelligence (AI) presents significant opportunities for parliaments to enhance their operations and to become more efficient and effective, enabling them to better serve citizens. However, adopting AI introduces new challenges and presents risks that must be carefully managed.

These *Guidelines for AI in parliaments* (the "Guidelines") have been developed for parliaments by parliamentary staff and the Inter-Parliamentary Union's (IPU) Centre for Innovation in Parliament (CIP). They provide comprehensive guidance to support parliaments on their journey towards understanding and implementing AI responsibly and effectively. By adopting a well-thought-through, strategic approach to AI, parliaments can harness the technology's full potential to drive innovation and efficiency in the legislative process.

The Guidelines cover key areas, including the potential role of AI in parliaments, related risks and challenges, suggested governance structures and AI strategy, ethical principles and risk management, training and capacity-building, and how to manage a portfolio of AI projects across parliament. They are complemented by a set of use cases, shared by parliaments, that describe how AI can support specific parliamentary actions.

The Guidelines stress the importance of a measured, risk-based approach to AI adoption. Key recommendations include the following:

- Start with small pilot projects to build experience.
- Focus on use cases with clear benefits and manageable risks.
- Ensure robust human oversight of AI systems.
- Prioritize transparency and accountability.
- Invest in data and AI literacy across the organization.
- Engage with diverse stakeholders throughout the process.

Audience

The Guidelines have been written to support a range of parliamentary roles:

- For **members of parliament (MPs)**, the Guidelines offer insights into the potential impact of AI on legislative processes, constituent engagement and parliamentary oversight. They provide a clear overview of AI capabilities and limitations, helping MPs to make informed decisions about AI adoption and regulation in parliament.

  For members serving on modernisation committees, committees of the future or similar bodies focused on technological advancements, the Guidelines provide strategic insights into AI governance and implementation. This is particularly valuable as parliamentarians increasingly face decisions about AI deployment within their own institutions while simultaneously developing legislation to govern AI use in society more broadly.

- For **senior parliamentary managers**, the Guidelines provide a high-level overview, offering strategic advice on developing AI governance frameworks, policies and oversight mechanisms. This includes establishing clear roles and responsibilities, creating codes of ethics, and aligning AI initiatives with parliament's organizational goals.

- For **staff involved in AI implementation**, the Guidelines provide detailed guidance on identifying use cases, managing projects, addressing technical challenges, upholding ethical standards and managing risks throughout the AI life cycle.

The Guidelines are designed to support parliaments of all sizes and levels of digital maturity – from large, well-resourced legislatures with advanced digital infrastructures, to smaller parliaments just beginning their digital transformation journey.

The Guidelines can be tailored to allow parliaments to focus on areas most relevant to their current needs and capabilities, and individual parliaments can adapt them to suit their unique circumstances, culture and resources. While digitally mature parliaments may be ready to implement more advanced AI applications, those at earlier stages can use the Guidelines to build foundational governance structures and develop AI literacy.

List of Guidelines

		Audience		
	Guideline	**For senior parliamentary managers**	**For MPs**	**For staff involved in AI implementation**
Key concepts	The role of AI in parliaments	✓	✓	✓
	Risks and challenges for parliaments	✓	✓	✓
	Alignment with national and international AI frameworks and standards	✓		✓
	Inter-parliamentary cooperation for AI	✓	✓	✓
Strategy	Strategic actions towards AI governance	✓		✓
	Generic risks and biases	✓		✓
	Ethical principles	✓		✓
	Introducing AI applications	✓	✓	✓
	Data and AI literacy	✓	✓	✓
Planning and implementation	Project portfolio management			✓

	Data governance			✓
	Security management			✓
	Risk management			✓
	Systems development			✓

About the Guidelines

Background

In April 2024, the Chamber of Deputies of Brazil – as the host of the CIP's Parliamentary Data Science Hub – convened, in Brasilia, a panel of experts from parliaments and the IPU to develop an outline of what would become the *Guidelines for AI in parliaments*. The intention was to develop comprehensive, flexible, scalable guidance to support parliaments as they embrace and responsibly engage with the rapidly emerging technologies associated with AI, whatever their level of digital maturity.

Acknowledgements

The CIP, the Parliamentary Data Science Hub, the IT Governance Hub and the editors are grateful for the contributions of the many parliamentary staff who helped shape, write and review these Guidelines, and for the support provided by the respective parliaments.

Editors

- Patricia Gomes Rêgo de Almeida, Chamber of Deputies of Brazil
- Ludovic Delépine, European Parliament
- Andy Williamson, Centre for Innovation in Parliament, Inter-Parliamentary Union

Contributors

- Patricia Rêgo de Almeida, Chamber of Deputies of Brazil
- Francisco Edmundo Andrade, Chamber of Deputies of Brazil
- Javier de Andrés Blasco, Chamber of Deputies of Spain
- Álvaro Carmo, National Assembly of Angola
- Virginia Carmona, Chamber of Deputies of Chile
- Giovanni Ciccone, Chamber of Deputies of Italy
- Ludovic Delépine, European Parliament
- Claudia di Andrea, Chamber of Deputies of Italy
- Michael Evraire, House of Commons of Canada
- Marcio Fonseca, Chamber of Deputies of Brazil
- José Andrés Jiménez Martín, Chamber of Deputies of Spain
- Vinicius de Morais, Chamber of Deputies of Brazil
- Rune Mortensen, Parliament of Norway
- Neemias Muachendo, National Assembly of Angola
- Jurgens Pieterse, Parliament of South Africa
- Manuel Pereira González, Senate of Spain
- Peter Reichstädter, Parliament of Austria
- Frode Rein, Parliament of Norway
- Esteban Sanchez, Chamber of Deputies of Chile
- Luciana Silo, Chamber of Deputies of Italy
- Paul Vaillancourt, House of Commons of Canada

- Kim van Dooren, Senate of Netherlands
- Marieke van Santen, Senate of Netherlands
- Rodolfo Vaz, Chamber of Deputies of Brazil
- Ricardo Vilarins, Chamber of Deputies of Brazil
- Andy Williamson, Centre for Innovation in Parliament, Inter-Parliamentary Union

Reviewers

- Patricia Rêgo de Almeida, Chamber of Deputies of Brazil
- Avinash Bikha, Centre for Innovation in Parliament, Inter-Parliamentary Union Ludovic Delépine, European Parliament
- Andy Richardson, Centre for Innovation in Parliament, Inter-Parliamentary Union
- Andy Williamson, Centre for Innovation in Parliament, Inter-Parliamentary Union

Graphic design

- Ana Paulla Diniz, Chamber of Deputies of Brazil
- Kelly Lima Cardoso, Chamber of Deputies of Brazil

About the Centre for Innovation in Parliament

The CIP was established in late 2018 to support parliaments in their digital transformation efforts. It facilitates knowledge-sharing and collaboration through a network of regional and thematic hubs, each hosted by a parliament or partner organization. The CIP has made significant contributions to parliamentary modernization by undertaking research and providing guidance on digital strategies, by organizing capacity-building activities, and by developing resources on topics such as open data, cybersecurity and emerging technologies, including AI. Through its work, including the *World e-Parliament Report* series and the *Guide to digital transformation in parliaments*, the CIP has become a key platform for parliaments worldwide to exchange good practices and innovative solutions, fostering more efficient, transparent and accessible legislative institutions.

Contact

For more information about this work, please contact innovation@ipu.org. We are always keen to learn how the Guidelines have been used. We welcome all feedback and suggestions.

Suggested attribution: Inter-Parliamentary Union (IPU), *Guidelines for AI in parliaments* (Geneva: IPU, 2024): www.ipu.org/AIguidelines.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Key concepts

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

The role of AI in parliaments

Audience

This high-level guideline is intended for parliamentary leadership and senior parliamentary managers, as well as for parliamentary staff and MPs who are interested in gaining a broad understanding of where AI can impact upon the work of parliaments.

About this guideline

This guideline offers an overview of parliamentary functions and suggests potential applications for AI within these functions. It explores how AI can transform various aspects of parliamentary work, from streamlining administrative tasks to enhancing legislative research and improving public engagement. It looks at how parliaments can leverage AI effectively, while emphasizing the importance of upholding democratic principles and values throughout the implementation process.

Potential uses for AI

There are opportunities to implement AI-based systems in many areas of parliament. AI can support administrative, legislative and public engagement processes, as well as parliamentary transparency efforts. While a small number of parliaments are already developing AI-based applications in some of the areas below, their usage in many other areas remains unexplored at this stage, or is only being tested on a pilot or prototype basis.

For examples of potential applications of AI in parliaments, refer to the accompanying use cases, which are intended to help parliaments develop their own proposals.

The first step in the journey is to understand where opportunities lie for a given parliament. It is important to recognize that every parliament is different and will want to seize the opportunities on offer differently, based on its own unique culture, working practices, resource availability and timing constraints. A framework to help identify opportunities and support the adoption of AI can be found in the following guidelines: Project portfolio management and Strategic actions towards AI governance.

Improving legislative research and analysis

AI systems can assist parliamentary staff in conducting comprehensive legislative research and analysis. Machine-learning algorithms can analyse vast volumes of legislative documents, identifying patterns, trends and relevant ideas. Additionally, AI-driven data analysis platforms can facilitate evidence-based policy formulation by synthesizing disparate sources of information and highlighting key findings for decision makers.

Analysing legislative trends

AI systems can analyse large volumes of legislative data (live or archived) in order to identify trends and patterns in proposed and enacted legislation. These insights enable MPs to better understand legislative priorities, areas of interest and areas requiring attention. They can also assist in identifying legislative key performance indicators (KPIs), helping parliamentary committees to determine the impact of legislation once enacted.

Identifying similarities and differences in legislation

Machine-learning algorithms can compare and contrast proposed draft legislation with legislation that already exists. Identifying similarities, disparities and areas of overlap (or contradiction) in this way can help to avoid conflicts between laws and ensure legislative coherence. This comparison also makes it possible to determine if two similar laws could have different impacts, and to identify the reasons for this.

Supporting the legislative function

AI systems can be used to assist the legislative work of parliament, supporting tasks such as assessing the impact of bills, and mapping amendments in order to understand their impact on a bill.

Analysing the impact of proposed legislation

AI systems can be used to analyse the potential impact of proposed legislation, helping MPs to make informed decisions about the viability and implications of a bill and to identify any unexpected consequences of the proposal.

Mapping bills and amendments

AI systems can be used to analyse amendments and voting to ensure that they are correctly matched up to bills, and that the changes (and the impact of those changes) are transparent and can be easily understood.

Assisting with post-legislative scrutiny

AI systems can be used to gather comprehensive data and provide a detailed analysis of the impact of legislation. This analysis enables MPs to understand whether a piece of legislation has met its aims and objectives, and to assess whether (and how) it might need revising.

Identifying stakeholders, experts and resources

AI systems can be used to analyse databases and online sources to identify experts, key stakeholders, research reports and other relevant resources related to specific legislative topics, ensuring that MPs' legislative work is supported by up-to-date and reliable information.

Analysing public submissions

The digital parliament is opening up opportunities for more people to submit their views to parliamentary committees and inquiries. While this is a positive development for democracy, it increases pressure on parliamentary staff, who have to manage and make sense of ever-larger volumes of qualitative data. AI systems can be used to assist with this task, and can be highly effective at identifying themes and grouping submissions.

Automating administrative tasks

AI technologies such as natural language processing (NLP) and robotic process automation (RPA) can automate routine administrative tasks such as scheduling meetings, drafting agendas and managing documentation. Offloading these repetitive tasks to AI systems allows more time and resources to be allocated to high-value activities, thus increasing productivity and efficiency.

Scheduling parliamentary meetings and sessions

AI systems can be used to analyse MPs' schedules, identify available slots, and automatically schedule meetings and parliamentary sessions, taking into account the availability of participants and meeting rooms and, therefore, avoiding coordination issues and inefficiencies.

Managing documents

Parliaments generate and handle large amounts of documents, such as bills, committee reports, communications and session minutes. AI systems can be used to automatically organize, label and archive these documents, facilitating quick search and retrieval when needed.

Automating translation

In multilingual parliaments, automatic document translation can be an invaluable tool. AI systems can be used to automatically translate legislative documents, allowing MPs to access information in their preferred language.

Managing digital communication

MPs receive large volumes of emails and other electronic communications from their constituents and colleagues. AI systems can be used to automatically classify these messages, identifying the most urgent or relevant ones and assigning them to the relevant person for processing.

Generating reports and analysing data

Report generation and data analysis are key supporting tasks in parliaments. AI systems can be used to collect and analyse relevant data, using data visualization

and predictive models to generate detailed and easily understandable reports for MPs.

Improving transparency

AI systems can play a pivotal role in promoting transparency and accountability within parliaments. Automated transcription and translation tools can generate accurate and timely transcripts of parliamentary debates and discussions, making legislative proceedings more accessible to the public, while AI-powered sentiment analysis tools can gauge public sentiment towards legislative proposals, enabling MPs to better understand and address their constituents' concerns.

Automating transcription of parliamentary debates

AI systems can be used to automatically transcribe parliamentary debates in real time. These accurate, rapidly produced transcripts can then be made available to the public and to specific users or departments within parliament, allowing citizens and key officials to access parliamentary proceedings without having to consult complete audiovisual recordings. Real-time speech-to-text transcription and translation – a possibility offered by some language models – could also facilitate effective communication in the context of large multinational events.

Visualizing legislative data

AI systems can be used to create interactive visualizations of legislative data, such as an MP's legislative activity or the progress of a bill through parliament. These visualizations make it easier for the public to understand and evaluate the work of parliament.

Accessing legislative information

AI-powered search tools can allow constituents to easily find information about bills, votes, committees and other aspects of parliamentary work. This promotes transparency by making legislative information more accessible and understandable to all.

Analysing economic data

AI systems can be used to analyse economic data related to parliamentary spending and the financial interests of MPs, helping to identify potential conflicts of interest, and promoting transparency and accountability by ensuring that MPs are subject to public scrutiny.

Producing plain-language summaries

AI systems can be used to summarize bills, reports and transcripts in plain language, making them easier to understand for ordinary citizens. Making such summaries available can enhance public participation in the legislative process and foster communication between MPs and their constituents.

Enhancing public engagement

AI-powered chatbots, sentiment analysis tools and multimedia content production systems can enhance public engagement with parliament by facilitating communication between citizens and MPs, offering insights into public opinion and helping to create accessible content.

Helping citizens connect with parliament

AI-powered chatbots and virtual assistants can enhance public engagement by providing fast, personalized responses to enquiries, facilitating communication between MPs and constituents, and disseminating information about parliamentary procedures and initiatives. These AI systems empower constituents to participate actively in the democratic process while relieving parliamentary staff of the burden of managing large volumes of enquiries manually.

Analysing sentiment in public spaces

AI-powered sentiment analysis tools can be used to monitor social media and other online platforms in order to assess public sentiment towards legislative topics, as well as towards MPs and their legislative decisions. These insights can help MPs understand their constituents' concerns and opinions, allowing them to better reflect these positions in parliament.

Producing multimedia content

AI systems can be used to automatically produce short video summaries that can then be posted on social media. These clips would typically focus on the most important part(s) of a speech or other intervention, with multilingual subtitles provided.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Risks and challenges for parliaments

Audience

This high-level guideline is intended for parliamentary leadership and senior parliamentary managers, as well as for parliamentary staff and MPs who are interested in gaining a broad understanding of the risks and challenges of AI.

About this guideline

This guideline examines the risks involved in the introduction of AI in parliaments, both from a strategic level and in terms of what it means operationally for parliaments, with a particular focus on risks that are unique to legislatures. For each of these areas, it provides brief checklists that senior parliamentary managers can use as a starting point to fully understand the impact that AI is likely to have.

As AI adoption gains traction and the use of this technology becomes more commonplace, parliaments must understand the implications and closely examine the risks and challenges associated with the implementation of AI.

Technologies such as generative AI, with its ability to create content based on vast amounts of data, promises productivity gains and potentially transformational change in parliamentary operations. However, it also introduces new complexities and risks that must be carefully managed.

For a discussion of the more generic risks and biases that AI introduces, refer to the guideline Generic risks and biases. For a discussion of the potential uses of AI in parliamentary settings, refer to the guideline The role of AI in parliaments.

Strategic considerations

At the strategic level, parliaments face several key challenges in adopting AI. Foremost among these is the development of comprehensive AI governance frameworks that ensure strong ethical principles, transparency and accountability in AI systems.

Parliaments must also address potential biases in AI algorithms, ensuring that these systems do not inadvertently amplify existing societal inequalities or underrepresent

minority views. This is particularly crucial in parliamentary contexts, where fair representation is a fundamental principle.

Public trust and perception present another strategic challenge. Parliaments must effectively communicate their use of AI to constituents, managing expectations and addressing concerns about the role of AI in democratic processes. This requires a delicate balance between showcasing the benefits of AI adoption and reassuring the public that human judgement remains central to parliamentary functions.

Strategic considerations checklist:

- Develop a comprehensive AI governance framework and policies that reflect parliament's ethical principles.
- Establish protocols for ensuring AI transparency and accountability.
- Expand the remit of existing data committees or similar bodies to encompass AI.
- Develop a communication strategy to inform the public about AI use in parliament.
- Regularly assess and mitigate potential biases in AI systems.

Unique parliamentary considerations

Several aspects of AI adoption are uniquely relevant to parliaments, with the rapid evolution of this technology requiring legislatures to develop flexible, future-proof AI strategies. This is particularly challenging given the typically slower pace of change in parliamentary institutions.

Unlike in other contexts, the successful implementation of AI in parliament requires buy-in across political divides. Different parties may have varying views on the role and extent of AI use in parliamentary functions, necessitating careful negotiation and compromise.

Perhaps most fundamentally, parliaments must balance the pursuit of efficiency through AI with the preservation of core democratic values. While AI can enhance many aspects of parliamentary work, it is crucial that it is used to augment – not replace – the essential human elements of democratic representation and decision-making.

The adoption of AI in parliaments has direct implications for legislative processes. There is the potential for AI to significantly alter how debates are conducted and how legislation is drafted. While AI can provide valuable insights and efficiencies, it is essential to maintain human oversight in any move towards AI-assisted lawmaking. The nuanced and often politically sensitive work of parliaments requires a level of judgement and ethical consideration that current AI systems cannot replicate.

Constituency engagement is another area where AI could have a profound impact. AI-powered tools could support deeper analysis of public sentiment on proposed legislation, potentially providing more real-time feedback. However, this must be balanced with the importance of direct constituent interactions to ensure that the human element of representation is not lost.

Unique parliamentary considerations checklist:

- Develop guidelines for maintaining human oversight in AI-assisted lawmaking and ensuring scrutiny of AI-based decisions.
- Establish rules for disclosing AI use in legislative processes.
- Develop a flexible AI strategy that can adapt with rapid technological change.
- Establish mechanisms for equitable access to AI resources across all of parliament.
- Develop protocols for international cooperation on AI in parliaments.

Operational challenges

On the operational front, implementing and integrating AI systems into existing parliamentary procedures and processes poses significant challenges, especially since these are often complex and steeped in tradition. Moreover, parliaments must ensure that AI adoption does not disrupt the essential human elements of political discourse and decision-making.

Data management and security are also critical concerns. Parliaments handle sensitive information and are prime targets for cyberattacks. AI systems could potentially create new vulnerabilities if they are not implemented with robust security measures.

Capacity-building and change management present another set of operational challenges. Developing AI literacy and data literacy among MPs and staff is crucial for effective use and oversight of these systems. However, this introduces a unique challenge in balancing the need for traditional parliamentary skills and knowledge with new AI competencies. Moreover, ensuring that AI systems are trained on high-quality data is essential to prevent biased or inaccurate outputs.

There is also the potential for job displacement within parliamentary staff, necessitating careful management of role redefinition and retraining.

Operational challenges checklist:

- Conduct a thorough assessment of existing parliamentary procedures for AI integration.
- Implement robust cybersecurity measures for AI systems.
- Develop and implement AI literacy and data literacy programmes for staff and MPs.
- Create a data quality assurance process for AI training data sets.
- Establish a change management plan to address potential job displacement and role changes.

Mitigation strategies

Given these challenges, a cautious and measured approach to AI adoption is advisable for most parliaments. The IPU's Centre for Innovation in Parliament recommends a step-by-step, risk-based approach.

Creating safe "lab environments" for AI experimentation is a prudent first step. This allows parliaments to explore potential use cases, such as producing summaries of

texts or creating records of debates, without risking core parliamentary functions. However, parliaments should be extremely cautious about introducing AI into core legislative systems at this stage.

Maintaining human scrutiny and control is paramount. Any AI-generated outputs must be explainable and subject to expert validation, and the entire system must be auditable. This is essential not only for ensuring accuracy but also for maintaining public trust in parliamentary processes.

Collaboration with other parliaments and external experts can be invaluable in building capacity and sharing best practices. The complexity of AI systems means that individual parliaments working alone may initially lack the skills and knowledge to implement AI-based systems safely and effectively.

AI policies and practices should be regularly reviewed and updated to account for the rapid pace of technological change. Parliaments must remain agile, continuously assessing the impact of AI on their operations and adjusting their approaches accordingly.

Mitigation strategies checklist:

- Create a safe environment for AI experimentation.
- Implement a step-by-step, risk-based approach to AI adoption and monitoring.
- Adopt and, where necessary, adapt these Guidelines to support the safe introduction of AI in parliament according to its organizational culture.
- Develop partnerships with other parliaments and external experts for knowledge-sharing.

Conclusion

The adoption of AI in parliaments offers significant potential benefits but also presents unique challenges. By carefully navigating the strategic, operational and legislative risks, parliaments can harness the power of AI to enhance their effectiveness while safeguarding the essential human elements of democratic representation.

Parliaments have a dual responsibility as both users and regulators of AI technology. They must lead by example in the responsible adoption of AI within their own institutions while also shaping the legislative frameworks that will govern AI use in broader society.

Parliaments therefore have an opportunity to set a standard for responsible AI use that could inform its adoption in other areas of government and in society at large. This positions parliaments at the forefront of defining how AI can be leveraged to strengthen – rather than undermine – democratic processes in the digital age.

The journey of AI adoption in parliaments is just beginning, and the path forward will require ongoing dialogue, rigorous oversight and a commitment to preserving the fundamental values of democratic governance.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Alignment with national and international AI frameworks and standards

Audience

This high-level guideline is intended for parliamentary leadership and senior parliamentary managers, as well as for parliamentary staff and MPs. It may also be useful for technical staff involved in the implementation of AI-based systems.

About this guideline

As parliaments consider adopting AI technologies, it is important to be aware of and, where appropriate, to adhere to government standards and relevant national and international frameworks for AI use. These standards and frameworks may significantly influence how parliaments implement AI and, in some cases, may require parliaments to adapt the approaches outlined in these Guidelines.

Compliance with AI regulations

Parliaments must take steps to ensure that their use of AI conforms to national laws, regulations and, where possible and appropriate, good practice for AI use in the public sector.

It should be noted that, while these Guidelines serve as an important starting point, parliaments may need to go above and beyond these in order to address their unique needs and uphold democratic principles. They should be used in conjunction with national and international standards to create a comprehensive approach to AI governance in parliament.

As part of its AI governance framework, parliament should stay informed about national and international AI-related policies and ensure that it remains compliant with these at all times. For example, the Chamber of Deputies of Italy has developed

a code of conduct for the use of generative AI, which emphasizes that the use of this technology must align with various multilevel strategies and regulations:

> [The Code of Conduct] for the Use of Generative Artificial Intelligence Tools has therefore been adopted, taking into account the Principles for the Use of AI in Support of Parliamentary Work, as laid down by the Supervisory Committee on Documentation Activities of the Chamber of Deputies, and having regard to the recommendations set forth in the 2024–2026 Three-Year Plan for ICT in the Public Administration, the Hiroshima Process International Code of Conduct for Advanced AI Systems as agreed by the G7, as well as the Guidelines for Secure AI System Development, promoted at the international level by the National Cyber Security Centre and signed on 27 November 2023 by the National Cybersecurity Agency.

In order to ensure compliance with applicable regulations, parliaments are advised to take the following steps:

- Research and identify relevant national and international AI frameworks or standards.
- Consult with government bodies responsible for AI oversight.
- Assess how these standards impact parliamentary AI implementation.
- Adjust internal AI policies and practices to ensure compliance.
- Regularly monitor national and international AI frameworks for updates.

Laws and regulations impacting upon the use of AI

In addition to AI-specific regulations, it is important to consider the implications of AI systems in terms of other laws and regulations that might be in place. These may cover the following matters, among others:

- Access to information and/or freedom of information
- Cybersecurity
- Data protection and privacy
- Discrimination and equality
- Employment
- Human rights and accessibility
- Intellectual property
- Procurement and competition

International standards and frameworks for AI

Parliaments may find the following international standards and frameworks helpful when defining their own governance processes, rules and regulations regarding the use of AI:

- *OECD Principles on Artificial Intelligence*: principles promoting innovative and trustworthy AI that respects human rights and democratic values
- *UNESCO Recommendation on the Ethics of Artificial Intelligence*: the first global standard-setting instrument on the ethics of AI
- *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*: guidance for ethical AI produced as part

of the Institute of Electrical and Electronics Engineers (IEEE) Global Initiative on Ethics of Autonomous and Intelligent Systems
- Standards by ISO/IEC JTC 1/SC 42: a list of standards developed by the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) joint technical committee on AI, including a data quality governance framework
- *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*: a framework convention developed by the Council of Europe's Committee on Artificial Intelligence
- *White Paper on Artificial Intelligence: A European approach to excellence and trust*: a European Commission document that, while specific to the European Union, provides a framework for trustworthy AI that could be informative for parliaments globally
- *G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI*: an OECD report that helped inform and structure discussions around the G7 Hiroshima Process
- *AI Governance: A Holistic Approach to Implement Ethics into AI*: a framework providing guidance on implementing ethical AI in practice
- *Risks, Harms and Benefits Assessment Tool*: a United Nations Global Pulse tool to help assess the risks and benefits of using AI in development and humanitarian contexts
- ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H): a joint International Telecommunication Union (ITU) and World Health Organization (WHO) focus group that has produced guidance and other outputs on AI governance in a critical public-sector domain

Parliaments should regularly check for updates to these frameworks and for new parliament-specific AI guidance documents as they emerge. The unique role of parliaments in democratic societies may necessitate the development of more tailored international guidance in the future.

National government frameworks and strategies for AI

For a list of national government frameworks, strategies and programmes relating to AI, refer to the sub-guideline Strategic actions towards AI governance: Find out more, which is part of the guideline Strategic actions towards AI governance.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in Parliaments

Inter-parliamentary cooperation for AI

Audience

This high-level guideline is intended for parliamentary leadership and senior parliamentary managers, as well as for senior IT staff who are interested in the adoption and application of AI in their parliament.

About this guideline

This guideline explores how parliaments can work together to support good practice and the ethical use of AI at a time when legislatures are increasingly embracing the opportunities and addressing the challenges presented by this technology. It looks at various forms of parliamentary cooperation and collaboration, providing guidance and recommendations for parliaments in this area.

Establishing inter-parliamentary AI networks

Every parliament will likely apply these Guidelines in unique ways, adapting them to its specific context, culture and needs. However, the universal nature of many AI challenges presents a compelling case for inter-parliamentary cooperation. Collaboration is encouraged at all levels, whether this is global, regional or subregional.

The foundation of effective collaboration lies in robust networks. Parliaments should create formal AI knowledge-sharing networks for continuous exchange. Networks such as the IPU's Centre for Innovation in Parliament (CIP) already exist as vehicles for these conversations, where parliamentary staff can discuss the latest developments and challenges in AI governance.

Sharing use cases, case studies and good practices

Parliaments are encouraged to document successful AI implementations, analyse challenges and share lessons learned. A centralized repository of use cases, case studies and solutions that is accessible to all parliaments – such as the use cases that accompany these Guidelines – can serve as a valuable resource for parliaments at various stages of AI adoption.

Collaborative AI project development

Joint research initiatives and shared pilot projects offer an opportunity for parliaments to pool resources and expertise, and to learn together. Parliaments could coordinate efforts, focusing on developing open-source AI tools specifically designed for parliamentary use. For instance, a collaborative project to create an AI-powered legislative drafting assistant could benefit many parliaments, including those that might not have the resources to undertake such an endeavour on their own.

Harmonizing AI governance frameworks

While every parliament operates within its unique context, there is value in aligning internal approaches to AI governance. Parliaments have already collaborated to create these Guidelines, and this process could be extended to grow the Guidelines and keep them up-to-date and relevant. Co-created resources such as these will provide a solid foundation for parliaments to build upon, adapting them to local needs while ensuring a baseline of good practice.

Capacity-building and training

The rapid evolution of AI necessitates ongoing training and capacity-building. Parliaments can create efficiencies by jointly developing training programmes for members and staff. They might even consider developing a certification in AI governance for parliaments. Exchanges of AI experts and specialists between parliaments will foster cross-pollination of ideas and expertise.

Data-sharing and standardization

Parliaments can jointly establish protocols for secure data-sharing and develop common data standards for AI applications, following the example of Akoma Ntosa, which was developed as an international standard for parliamentary data. This could include collaborative efforts to improve data quality and consistency across parliaments, enhancing the potential for AI-driven insights.

Ethical AI evaluation and auditing

Parliaments can work together to develop shared frameworks for AI system audits, facilitating peer-review processes for AI implementations. Together, they can develop common metrics for measuring AI impact and effectiveness, ensuring that AI serves the needs of democratic institutions and citizens.

Addressing common challenges

Parliaments could coordinate on collaborative approaches to common issues such as mitigating biases, ensuring AI transparency or managing AI-related privacy concerns. By pooling knowledge and resources, parliaments could develop more effective solutions to these shared challenges.

Engaging with international initiatives

As AI governance becomes a global concern, parliaments should engage with international initiatives to ensure their interests are represented in global forums. By taking these steps, parliaments can facilitate collaboration on AI-related legislation and regulation, thus helping to ensure that democratic principles are upheld in the global AI landscape.

Future-proofing AI governance

Parliaments should be involved in collaborative foresight and scenario planning exercises, helping to prepare for potential future developments in AI. This could include joint research on emerging AI technologies and their implications for parliamentary work, as well as the development of adaptive governance frameworks that can evolve alongside AI technology.

The role of the IPU's Centre for Innovation in Parliament

Throughout these collaborative efforts, the CIP acts as a crucial facilitator and coordinator, uniquely positioned to perform the following roles:

- Acting as a central hub for the sharing of knowledge and good practices, and supporting regional networks
- Providing a neutral platform for discussing challenges and developing solutions
- Representing parliamentary interests in global AI governance discussions
- Fostering a community of practice among parliaments, and encouraging ongoing dialogue and collaboration
- Offering expertise and resources to support parliaments at various stages of AI adoption

By leveraging its network and resources, the CIP plays a pivotal role in ensuring that parliaments worldwide are well-equipped to harness the benefits of AI while mitigating its risks.

Conclusion

As AI continues to transform parliamentary work, collaboration becomes not just beneficial, but essential. Through shared efforts in areas such as exchanging knowledge, developing projects and governance frameworks, and addressing common challenges, parliaments can navigate the complex landscape of AI more effectively.

The path forward is clear: through collaboration and coordination, parliaments can lead the way in ethical and effective AI governance, setting a standard for responsible AI use that extends to other sectors.

Actions

- Join the CIP's thematic and regional networks to get peer support.
- Document and share AI use cases and implementation case studies.

- Collaborate on developing common AI governance frameworks, training programmes and data standards for parliaments.
- Engage in international AI initiatives and forums in order to represent parliamentary interests in global AI governance discussions.
- Participate in or initiate joint AI projects, focusing on open-source tools specific to parliamentary needs.
- Conduct collaborative foresight exercises to prepare for future AI developments and their implications for parliamentary work.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Strategy

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Strategic actions towards AI governance

Audience

This high-level guideline is intended for senior parliamentary managers, including those working at a strategic or board management level, and in particular those who have, or will have, a role in governing the use of AI.

About this guideline

This guideline outlines key actions for effective AI governance in parliaments. It focuses on creating an ethical framework, establishing robust governance structures and developing comprehensive policies for the use of AI-based technologies. It emphasizes stakeholder engagement, strategic alignment with parliamentary goals, and capacity-building.

This guideline promotes responsible AI innovation, balancing technological advancement with ethical use to deliver benefits including improved data reliability and risk management, and more effective AI implementations.

Individual parliaments should adapt the recommendations contained in this guideline to suit their unique circumstances and resources, using them as a foundation for developing tailored guidance aligned with their existing strategies and methods.

Why strategic AI governance matters

The journey towards effective AI governance begins with an understanding of its importance. AI governance is about more than simply managing new technology: it is about creating a framework that maximizes the benefits of AI while minimizing its risks.

By taking a strategic approach to AI governance, parliaments can build a robust AI ecosystem that balances operational needs with innovation and ethical considerations, enhancing data and system reliability, and creating a trustworthy ecosystem for AI-based digital services.

Strategic governance ensures that AI initiatives align closely with business requirements, increasing the overall effectiveness of AI implementations. It promotes

fair and inclusive practices, upholding the democratic principles that parliaments embody. By encouraging responsible innovation, it positions parliaments at the forefront of technological advancement in governance.

Moreover, a strategic governance approach improves compliance with still-evolving legislation governing AI use in the public sector, helping parliaments to stay ahead of the curve and reducing the risk of future non-compliance.

Key benefits of strategic AI governance:

- Maximizes the benefits to be obtained from AI while better managing risk
- Ensures clear alignment with business processes
- Increases the reliability of data and systems
- Promotes fair practices and responsible innovation
- Increases parliament's preparedness for evolving regulatory landscapes
- Leaves parliaments well-placed to cooperate and collaborate, sharing good practices relating to AI

Governance structure and policy

The first step in this journey is to establish a solid governance structure, which requires a multidisciplinary approach across parliament involving executive boards, legal departments, business units and the IT function. Some parliaments may choose to add AI capacity to existing governance boards, while others might create an AI-specific board. Regardless of the chosen structure, the responsibilities of such a body are to approve AI policies, monitor strategies, oversee projects and address ethical issues as they arise.

With this structure in place, the next task is to develop a comprehensive AI governance policy. This policy should guide all AI-related activities within parliament. A dedicated working group should lead this effort, defining objectives, outlining ethical principles and ensuring legal compliance. The policy will clearly delineate roles and responsibilities for key stakeholders, as well as establish processes for risk management, information security and data governance.

Actions:

- Establish an AI governance structure and policy
- Define roles, responsibilities and key processes

Ethical and responsible use of AI

Since the adoption of AI is both challenging and filled with potential, adopting these strategic measures will help parliaments lead the way in responsible AI use. It is possible to leverage AI's immense potential while upholding democratic principles and ethical standards.

As parliaments embark on this journey, they must do so not just for their own benefit, but also for the benefit of the citizens they serve. By governing AI responsibly, they can pave the way to a future where technology and democracy work hand in hand, enhancing governance and improving lives.

Key points:

- Responsible AI governance benefits both parliaments and citizens.
- Parliaments can lead the way in responsible AI use.

Ethical foundations

At the heart of AI governance lies a robust code of ethics. This code serves as a declaration of values, guiding the use of AI and the management of the associated risks. It should reflect parliament's commitment to privacy, transparency, accountability, fairness and societal well-being. Importantly, this code must align with existing laws, regulations and parliamentary procedures. It must be adaptable and responsive to change. It should also include real-world examples, in order to aid understanding and adoption and to demonstrate how these ethical principles apply in practice.

Actions:

- Create an AI code of ethics reflecting parliamentary values.
- Ensure alignment with national laws, standards, international good practice and parliamentary procedures.

AI strategy and capacity-building

With the foundational elements in place, parliament can now focus on creating a comprehensive AI strategy. This strategy serves as a road map, aligning AI use with the institution's broader goals and objectives. It should include a clear vision, set measurable goals, outline specific actions and define key performance indicators (KPIs). The strategy must also address ethical principles, regulatory issues and infrastructure requirements.

Once the decision to adopt AI has been made, building capacity within parliament becomes crucial for success. This involves developing a robust plan for staff training and skills development.

Thinking about how to approach the introduction of AI is also important at this stage. It is often beneficial to start with small, manageable pilot projects, which help to build confidence and demonstrate the value of AI in a controlled environment. Encouraging experimentation and cross-functional collaboration can foster innovation and drive successful AI initiatives.

Actions:

- Develop a comprehensive AI strategy aligned with parliamentary goals.
- Build internal capacity through training and skills development.
- Look for small pilot projects to build confidence and demonstrate value.

Stakeholder engagement

Even the best governance structure and code of ethics will prove ineffective without proper stakeholder engagement. Identifying and involving stakeholders from various levels and departments is crucial. Engaging these stakeholders early and often helps

to manage risks, identify potential pitfalls and significantly increase the chances of successful AI implementation.

Action:

- Engage stakeholders from all relevant departments and levels.

Find out more

- For a list of national government guidelines, strategies and programmes relating to AI, refer to the sub-guideline Strategic actions towards AI governance: Find out more.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Strategic actions towards AI governance: Policy and structure

About this sub-guideline

This sub-guideline is part of the guideline Strategic actions towards AI governance. Refer to the main guideline for context and an overview.

This sub-guideline provides guidance and recommendations for the design and development of AI governance policies and structures in parliaments.

Background

When a parliament decides to adopt AI-based systems and services, it embarks on a journey that requires careful planning and a multidisciplinary approach. This journey begins with the recognition that AI governance is not just an IT issue: it is a matter that touches every aspect of parliamentary operations.

The first step is to assemble a diverse team. Executive boards, legal departments, and business and IT units all have crucial roles to play. This team will work together to create a governance structure that integrates business needs, legal and regulatory considerations, and technological insights. The exact nature of this structure will vary from parliament to parliament, reflecting each institution's unique culture and existing working methods.

Developing an AI governance structure

When it comes to establishing a governance structure, parliaments have two main options: either they add AI capacity to an existing board, or they create a dedicated AI-specific board. Both approaches have their merits, and the choice will depend on each parliament's specific circumstances and resources.
Regardless of the chosen approach, the AI governance structure will have a wide-ranging set of responsibilities, including the following:

- Approving and monitoring AI policy and strategy
- Overseeing the development of an AI code of ethics

- Managing budgets for AI research and development
- Supervising AI projects from initiation to completion
- Dealing with ethical issues as they arise
- Monitoring AI systems throughout their life cycle, including defining criteria for when these systems should be decommissioned

Within this framework, it is crucial to define key bodies. These typically include the following:

- A high-level “governing body” (which may be part of IT or corporate governance)
- A body responsible for IT and data science
- Business units that act as stakeholders in the AI system life cycle
- An ethics committee – either a new entity or an adapted existing data committee – to manage the unique ethical challenges posed by AI

While a central group should oversee these roles, day-to-day responsibilities can be distributed across various areas of parliament. This might involve assigning tasks to existing functional areas or creating new units specifically to manage AI-related work. This approach ensures comprehensive governance while allowing for flexibility in implementation.

By taking these steps, parliaments can create a robust governance structure that enables them to harness the benefits of AI while effectively managing its risks and ethical implications.

Actions:

- Assemble a multidisciplinary team from across parliament to lead AI governance efforts.
- Choose and implement either an integrated or a dedicated AI governance board structure.
- Define and assign key roles and responsibilities for AI governance, including policy approval, ethical oversight and project management.
- Establish or adapt an ethics committee to address AI-specific ethical challenges.
- Develop a clear AI life cycle management process, from project initiation to system decommissioning.

Establishing an AI governance policy

The purpose of a parliamentary AI governance policy is to establish a unified approach to AI use within the parliamentary environment. This policy will set measurable standards and provide top-level monitoring for AI implementation.

The policy development process involves creating a dedicated working group, defining objectives, and establishing ethical principles aligned with national regulations and international best practices. It must address legal compliance, data protection and digital public service legislation.

The policy will outline roles and responsibilities for all stakeholders involved in the AI life cycle, from data scientists to legal experts. It will also define AI-supported business processes, prioritizing them while considering risk tolerance and regulatory requirements. This includes specifying conditions for AI use, prohibited areas, and approval processes for certain AI applications.

Guidelines for AI sourcing, development and outsourcing will be established, along with clear communication strategies to inform all parliamentary staff and MPs about the policy.

Regarding generative AI, the policy will provide clear usage guidelines and detail necessary precautions. While recognizing potential risks, it will also encourage innovative experimentation in a controlled manner, avoiding outright bans that might lead to unauthorized use on personal devices.

Actions:

- Establish an AI policy working group to lead the development process.
- Outline roles, responsibilities and processes for AI governance and implementation.
- Create guidelines for AI use, including prohibited areas and approval processes.
- Define mechanisms for ensuring regulatory compliance.
- Develop a communication strategy to inform all staff and MPs about the AI policy.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Strategic actions towards AI governance: Strategy and innovation

About this sub-guideline

This sub-guideline is part of the guideline Strategic actions towards AI governance. Refer to the main guideline for context and an overview.

This sub-guideline provides guidance and recommendations on AI strategy and innovation for parliaments.

Background

The journey towards effective AI governance begins with an understanding of its importance. AI governance is about more than simply managing new technology: it is about creating a framework that maximizes the benefits of AI while minimizing its risks.

By taking a strategic approach to AI governance, parliaments can build a robust AI ecosystem that balances operational needs with innovation and ethical considerations.

Key points:

- AI governance is crucial for maximizing benefits and minimizing risks.
- A strategic approach balances operational needs, innovation and ethics.

Creating an AI strategy

Once the foundations are established in terms of strong governance and an AI code of ethics, and once robust engagement with key stakeholders has identified where AI can add value to the work of parliament, it is time to turn to developing an AI strategy grounded in this work.

An AI strategy is a comprehensive plan that outlines how parliament will use AI to achieve its goals and address challenges. It can be part of parliament's broader

strategy or a specific document that is aligned with this strategy. It is a corporate-level strategy that has goals which depend on AI, not a technology strategy. The main aim is to develop a road map for AI solutions that is aligned with business needs.

Focusing on parliament's goals, an AI strategy encompasses AI systems and the requirements for developing, deploying or purchasing them, with consideration given to ethical principles, as well as to regulatory and legislative issues. Base requirements to drive the strategy forward include workforce planning and infrastructure.

As a high-level document prepared for, and agreed by, senior parliamentary leaders, an AI strategy should use business language and arguments for business decision makers. Many senior managers will therefore already be familiar with its format and structure, and parliament should adopt a structure and approach it already uses, if appropriate. Alternatively, parliament can follow the sample structure outlined below and illustrated in Figure 1 (below):

Figure 1: Examples of goals, actions and KPIs in a parliamentary AI strategy

Vision

Formulate a clear vision statement indicating what parliament's needs are for AI. The nature of this statement will depend on whether the AI strategy is a stand-alone document specific to AI, or if it is integrated into a broader parliamentary strategy. In the latter case, there should be a single, overarching vision.

Goals

Include measurable goals that parliament can achieve using AI systems. These goals can focus on processes, practices and resources aimed at improving or driving AI adoption or mitigating AI-related risks.

Actions or drivers to achieve the goals

Detail specific projects and initiatives that will be implemented to achieve the stated vision and goals. These projects can affect multiple goals at once. Likewise, a single goal can be impacted by multiple projects. Projects can directly address business needs, prerequisites for business needs, or processes and practices to mitigate risks.

Key performance indicators (KPIs)

State what metrics will be used to measure progress towards the goals. It is often useful to set a target for each KPI.

Risks

Identify the main risks associated with the inappropriate use of AI, which can justify specific actions within the strategy.

Adopting an agile approach

An agile approach enables rapid iteration and continuous improvement, making it a practical way to work when innovating with new technologies such as AI:

- Regularly reviewing and adjusting AI projects based on feedback and changing business needs is important.
- Highlighting quick wins and sharing successes helps to build momentum and show the value of AI.
- Publicizing early successes and lessons learned encourages wider adoption across the organization.
- Engaging leadership by securing agreement from top executives and aligning AI initiatives with strategic business goals ensures ongoing support and commitment.
- Keeping leadership informed and involved in AI projects is crucial.

Managing change

Adopting rigorous change management practices within the iterative development process helps parliaments to manage resistance and ensure the smooth adoption of AI technologies. It is essential to develop a clear change management plan and to transparently communicate the goals of AI adoption, as well as the technology's impact on the organization, its staff and members. By understanding and carefully navigating the traditionally conservative culture of parliament and demonstrating clear, tangible benefits from the adoption of AI, it is possible to foster innovation and drive successful AI initiatives.

Promoting innovation

AI, as a new technology with immense potential, is very much about innovation. A good AI governance regime will include ways to promote innovative practices, taking a strategic and nuanced approach.

The first step is to build a strong case for AI by clearly articulating its benefits and focusing on how it can address specific business needs and challenges.

Demonstrating successful AI implementations in similar organizations through data and case studies can be persuasive. Likewise, selecting projects with a high likelihood of success can build confidence and show the value of AI.

Starting with small, manageable pilot projects that have clear objectives and measurable outcomes is crucial. This approach builds knowledge and experience, helps to develop familiarity and trust in AI-based systems, and can demonstrate potential, serving as a catalyst for further innovation. Of course, because pilots are also about experimenting and testing ideas, it is important to accept that some will inevitably fail. In other cases, it may be determined that the pilot is not worth pursuing. Building a reflective learning process into the innovation cycle will help parliaments to realize value and learn lessons as they go.
Innovation can be supported through the following approaches:

- Planning for education and awareness-raising
- Building cross-functional teams
- Encouraging a culture of experimentation
- Leveraging external expertise and partnerships

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Strategic actions towards AI governance: Stakeholder engagement

About this sub-guideline

This sub-guideline is part of the guideline Strategic actions towards AI governance. Refer to the main guideline for context and an overview.

This sub-guideline provides guidance and recommendations on engaging with stakeholders from across parliament as part of the AI governance and adoption process.

Identifying and engaging with internal stakeholders

Establishing good AI governance in parliament involves many stakeholders from across the institution. These will come from a range of levels and from across many different functional areas or organizational units.

Building strong engagement is crucial for creating buy-in within parliament as well for identifying risks, challenges and opportunities. Identifying which stakeholders need to be engaged with and then building that engagement with them early in the process of AI adoption is vital, helping to build support, knowledge and understanding across parliament.

Since the organizational structure and size of parliaments varies greatly, certain roles may not exist in specific parliaments (especially in smaller or less well-resourced legislatures). However, as a general rule, parliament should consider engaging on the following matters with the bodies, units or teams listed below:

- Approve AI policy, strategy and budget: Senior leadership
- Develop and support AI policy: Legal staff
- Identify legal risks and implications: Legal staff
- Identify areas for AI value addition: Business staff
- Implement data literacy and AI literacy programmes: Training staff
- Ensure practical grounding of strategic actions: IT and data staff
- Manage AI implementation risks: Risk management team

- Ensure compliance with regulations: Compliance team
- Address ethical considerations: Ethics and data committee
- Plan financially for AI projects: Finance department
- Manage communications about AI initiatives: Communications team
- Oversee AI project implementation: Project management office

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Strategic actions towards AI governance: Find out more

About this sub-guideline

This sub-guideline is part of the guideline Strategic actions towards AI governance. Refer to the main guideline for context and an overview.

National government guidelines, strategies and programmes

Below is a list of national government AI guidelines, strategies and programmes relating to AI:

- **Argentina**: Presidencia de la Nación: Plan Nacional de Inteligencia Artificial (available in Spanish only)
- **Australia**: Australian Government: Australia's Artificial Intelligence Action Plan
- **Brazil**: Chamber of Deputies: Digital Transformation Strategy of Brazilian Chamber of Deputies 2021–2024
- **Brazil**: Ministry of Science, Technology and Innovation: Summary of the Brazilian Artificial Intelligence Strategy (EBIA) 2021
- **Canada**: Government of Canada: Pan-Canadian Artificial Intelligence Strategy
- **Denmark**: Agency for Digital Government: The Danish National Strategy for Artificial Intelligence.
- **Germany**: German Federal Government: Artificial Intelligence Strategy of the German Federal Government – 2020 Update
- **Hungary**: Ministry for Innovation and Technology: Hungary's Artificial Intelligence Strategy 2020–2030
- **Italy**: Italian Artificial Intelligence Strategy: Italian Strategy for Artificial Intelligence 2024–2026
- **Japan**: Strategic Council for AI Technology: Artificial Intelligence Technology Strategy

- **Norway**: Ministry of Digitalisation and Public Governance: The National Strategy for Artificial Intelligence
- **Spain**: Government of Spain: National Artificial Intelligence Strategy (ENIA)
- **Switzerland**: Federal Council: Guidelines on Artificial Intelligence for the Confederation: General frame of reference on the use of artificial intelligence within the Federal Administration
- **United Kingdom**: Government of the United Kingdom: National AI Strategy

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Generic risks and biases

Audience

This high-level guideline is intended for senior parliamentary managers, as well as for parliamentary staff and MPs interested in gaining a broad understanding of generic risk and biases associated with AI.

About this guideline

This guideline describes a range of generic risks and biases related to the implementation of AI technologies, which parliaments will need to understand before embarking on AI projects and initiatives.

For a discussion of risks that relate more specifically to the unique work of parliaments, refer to the guideline Risks and challenges for parliaments.

Why inappropriate AI use is a risk

Inappropriate AI use can entail risks at various levels, from the individual to the global:

- Unintended consequences, such as reinforcing existing biases through biased systems, resulting in unfair treatment of individuals or groups
- A lack of accountability and transparency – which are crucial for building user trust – owing to poor understanding of the complexity of AI systems and the underlying decision-making processes
- The manipulation of public opinion through deepfakes, misinformation and automated propaganda
- The creation of echo chambers, which amplify biased views and extremism
- Psychological profiling, which allows for the targeted manipulation of individuals
- Fake content, which can potentially contribute to eroding trust in genuine information
- Behavioural nudging, which can subtly influence opinions and actions, often without people's full awareness, thus potentially undermining democratic processes and informed discourse
- Physical and psychological harm through the use of AI systems in health care, autonomous vehicles and industrial automation, which can lead to accidents or malfunctions

- Issues such as addiction, anxiety and depression, caused by AI-driven social media algorithms that promote harmful content or create unrealistic social comparisons
- Stress, privacy invasion and discrimination through the use of AI systems for surveillance and profiling, exacerbating mental health problems and social tensions

Categories of risk

The integration of AI introduces new types of risk that may not be familiar to parliaments. These can include the following:

- Lack of AI literacy
- Bias and discrimination
- Privacy invasion
- Security vulnerabilities
- Lack of accountability
- Job displacement
- Ethical dilemmas
- Shadow AI
- Lack of data sovereignty
- Lack of trust

For further discussion of these categories, refer to the sub-guideline Generic risks and biases: Categories of risk.

Identifying biases in a parliament

Bias is a systematic difference in the treatment of objects, people or groups compared to others, leading to an imbalance in the distribution of data.

Biases are part of people's lives. They usually start with habits or unconscious actions (cognitive biases) which, over time, materialize as technical biases (data biases and processing biases). Such a scenario increases or creates risks that could result in untrustful AI systems.

Biases in AI systems arise from human cognitive biases, the characteristics of the data used or the algorithms themselves. Where AI systems are trained on real-world data, there is the possibility that models can learn from, or even amplify, existing biases.

In a statistical context, errors in predictive systems are the difference between the values predicted as model output and the real value of the variables considered in the sample. When the error occurs systematically in one direction or for a subset of data, bias can be identified in the data treatment.

Cognitive biases

Cognitive biases are systematic errors in judgements or decisions common to human beings owing to cognitive limitations, motivational factors and adaptations

accumulated throughout life. Sometimes, actions that reveal cognitive biases are unconscious.

For a list of cognitive biases, refer to the sub-guideline Generic risks and biases: Cognitive bias types.

Data biases

Data biases are a type of error in which certain elements of a data set are more heavily weighted or represented than others, painting an inaccurate picture of the population. A biased data set does not accurately represent a model's use case, resulting in skewed outcomes, low accuracy levels and analytical errors.

For a list of cognitive biases, refer to the sub-guideline Generic risks and biases: Data bias types.

Processing and validation biases

Processing and validation biases arise from systematic actions and can occur in the absence of prejudice, partiality or discriminatory intent. In AI systems, these biases are present in algorithmic processes used in the development of AI applications.

For a list of cognitive biases, refer to the sub-guideline Generic risks and biases: Processing and validation bias types.

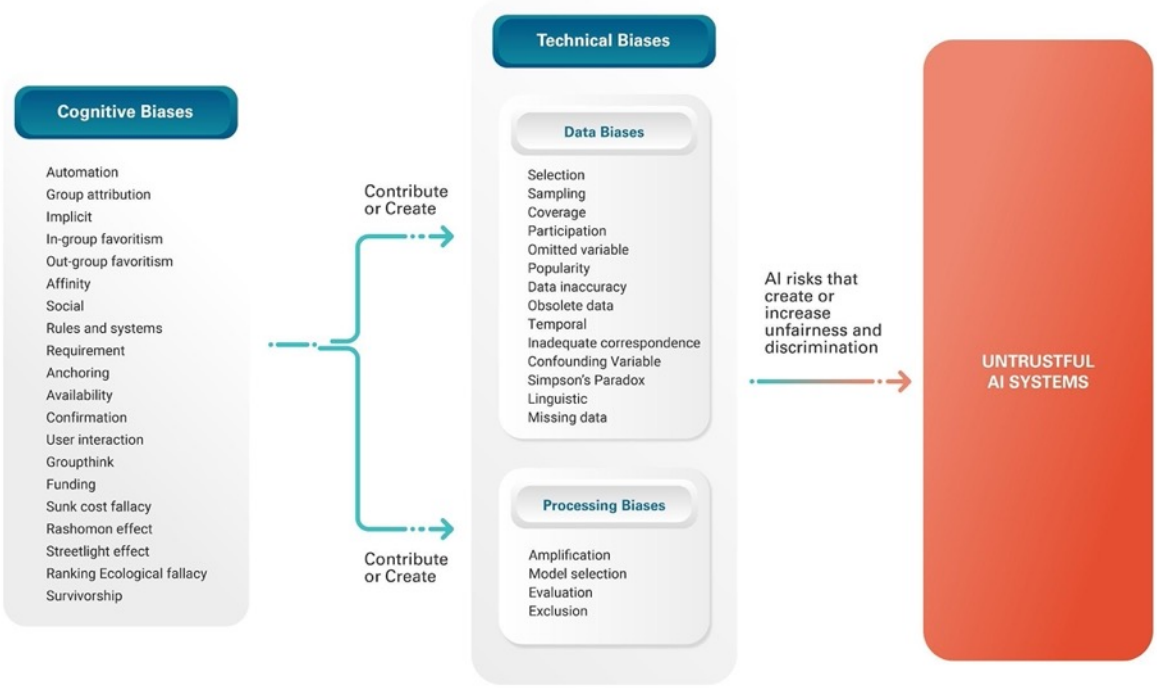
Interrelationship between biases

Cognitive biases are part the culture of many societies and organizations. They are often present, unconsciously, in the work processes and decisions that underpin the functioning of institutions. Over the years, cognitive biases are transformed – often in combination – into data biases and processing biases.

The underrepresentation or omission of a particular type of data in a data sample can therefore be the result of one or more of the following factors (among others):

- Systems were built by teams that unconsciously did not involve other organizational units owing to incorrect judgements regarding their participation.
- Important stakeholders were not involved in the design of data-entry systems because they had a different view than project managers.
- System interfaces favoured individual points of view or confirmed preconceived ideas.
- Irrelevant or incomplete databases were used to train AI systems simply because they were easy to obtain and avoided the need for negotiation between managers from different departments.
- AI system projects that revealed decisions based on inappropriate variables were launched anyway in order to justify the costs already incurred.
- AI system developers were so used to working with certain models that they used them in situations where they were inappropriate.

Figure 1: Bias path from the unconscious to untrustful AI systems



Source: Adapted from [NIST Special Publication 1270](#) and the [Oxford Catalogue of Bias](#)

Some biases can multiply the impact of others

Below are some examples of how cognitive biases can influence and, in some cases, even compound data or processing biases in parliamentary settings:

- Parliament feeds data sets with information from surveys and questionnaires completed only by people sharing the same political party ideology. Here, there is a high likelihood of existing affinity bias. Moreover, if this data set contains data such as “opinion regarding a specific theme”, and it is used to train an AI algorithm, there is a high possibility that such biases could be reproduced in that AI system.
- Parliament uses only data sets from a very small number of committee meetings to train an AI algorithm. In this case, there is a likelihood of interpretation biases because some terms may have different meanings or importance to different committees.
- Parliament has spent its entire innovation budget but the project team has failed to find the best AI algorithm to solve the original problem. The team implements an AI system anyway, launching it as a successful innovation, in an attempt to justify the costs. This is a funding bias that results in an AI system that is not reliable.

As the examples below show, with generative AI tools, all cognitive biases contained in a vast data set can be combined together and exposed directly to the user:

- A generative AI tool replicates bias against female job applicants when asked to draft letters of recommendation. Letters for male applicants often use terms like “expert” and “integrity” while female candidates are described using terms such as “beauty” and “delight”.
- Male researchers using a generative AI tool to create avatars receive diverse, empowering images portraying them as astronauts and inventors. However, a female researcher receives sexualized avatars – including topless versions, reminiscent of anime or video-game characters – that she did not request or consent to.
- A generative AI system fails to create appropriate images of people with disabilities.

Find out more

- The State of Data & AI Literacy Report 2024
- Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World
- Crowdsourcing Moral Machines
- Embedded ethics: some technical and ethical challenges
- UNESCO Recommendation on the Ethics of Artificial Intelligence
- Oxford Catalogue of Bias
- NIST Special Publication 1270: Towards a Standard for Identifying and Managing Bias in Artificial Intelligence
- ISO/IEC TR 24027:2021: Information technology – Artificial intelligence (AI) – Bias in AI systems and AI aided decision making

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Generic risks and biases: Categories of risk

About this sub-guideline

This sub-guideline is part of the guideline Generic risks and biases. Refer to the main guideline for context and an overview. For a discussion of risks that relate more specifically to the unique work of parliaments, refer to the guideline Risks and challenges for parliaments.

This sub-guideline explores new types of risk arising from the integration of AI that may not be familiar to parliaments and that, if not addressed effectively, can undermine democratic processes and public trust in parliamentary institutions.

Lack of AI literacy

AI literacy is an understanding of the basic principles, capabilities and limitations of AI – something that is crucial for informed decision-making about AI adoption and oversight in parliaments. It involves the ability to recognize AI applications, grasp fundamental concepts like machine learning and data analysis, and critically evaluate AI's potential impacts. Without adequate AI literacy, users may misinterpret AI results, fail to recognize discriminatory patterns, become overly reliant on flawed AI systems, and overlook ethical and legal implications. This can lead to poor decision-making and potential harm.

Bias and discrimination

AI systems used in parliamentary functions, such as for automated decision-making or policy analysis, can reflect and reinforce cognitive and other biases present in their training data. This can result in skewed policy recommendations and discriminatory legislative outcomes, adversely affecting minority groups, and undermining the principles of equality and fairness that underpin democratic institutions.

Privacy invasion

Parliamentary systems often handle sensitive personal and political data. Improper data-protection measures can lead to privacy infringements when using AI for data analysis and decision-making. Unauthorized access to, or misuse of, this data can

compromise the privacy of citizens, MPs and other stakeholders, eroding trust in parliamentary processes.

Security vulnerabilities

AI systems, particularly those used in parliamentary settings, are potential targets for cyberattacks. These attacks can lead to the manipulation or theft of sensitive legislative data, but can also disrupt parliamentary operations or compromise the integrity of legislative processes. This poses significant risks to national security and public safety.

Lack of accountability

The opaque nature of AI decision-making – often termed the “black box” problem – presents challenges in parliamentary contexts where transparency and accountability are paramount. Decisions made or influenced by AI without clear explanations can lead to difficulties in holding the right entities to account for legislative outcomes, diminishing public trust in democratic institutions.

Job displacement

While AI can improve efficiency, the automation of administrative tasks within parliamentary functions can lead to job and task displacement, particularly for support and administrative staff. As AI becomes increasingly adept at handling routine tasks such as scheduling, document processing and data analysis, the need for human involvement in these roles may decrease. This reduction in demand can lead to workforce downsizing, resulting in unemployment and economic disruption for those affected.

Aside from the loss of jobs, the nature of remaining roles may change significantly. Tasks that were once performed by human workers may be automated, leading to a shift towards more complex, decision-oriented or creative responsibilities that require a higher level of expertise. This evolution in job tasks can be challenging for employees who may not have the skills or experience needed to adapt, creating further risks of job insecurity and potential displacement.

The shift towards AI-driven processes also has the potential to increase job polarization, where low-skill, routine jobs are automated, leaving a gap that may not easily be filled by existing employees. This could exacerbate social and economic inequalities, particularly if the affected workers are unable to transition into new roles that require different skills.

Ethical dilemmas

AI applications in parliamentary settings raise ethical questions, particularly regarding the delegation of decision-making authority. Relying on AI for policy recommendations, legislative drafting or constituent services can lead to ethical dilemmas, especially if AI decisions conflict with human values or lack the necessary contextual understanding. Different AI services may report varying values depending on the country in which the underlying model is defined and trained.

Shadow AI

Shadow AI, which is related to the concept of shadow IT, can be defined as the unsupervised or unsanctioned use of generative AI tools within an organization or institution outside of its IT and cybersecurity framework. Shadow AI can expose organizations to the same risks as shadow IT: data breaches, data loss, non-compliance with privacy and data protection regulations, lack of oversight from IT governance, misallocation of resources, and even new risks stemming from a lack of understanding of the technology, such as the creation of AI models with biased data that can produce incorrect results.

Lack of data sovereignty

Training and deploying AI systems demands massive computing and storage resources, often requiring the use of public cloud systems. In some cases, these cloud systems are located in a different country and are therefore subject to the laws and regulations of that country. Without appropriate risk-mitigation strategies, such as encryption or data minimization, it may be difficult for parliaments to maintain effective control over such AI systems.

Lack of trust

The adoption of AI systems in parliamentary functions carries significant risks related to a lack of trust. One of the primary concerns is the complexity and opacity of these systems, which can lead to uncertainty about whether they are providing accurate and reliable information.

The absence of clear information on how these systems respect privacy or the nature of the data used for training further exacerbates distrust. Users may be concerned that their data could be misused or that the AI system's decisions are biased or flawed owing to inadequate or biased training data. This lack of trust can hinder the effective integration of AI in parliamentary operations, as stakeholders may be reluctant to rely on systems they do not fully understand or trust.

The overall risk is that without trust, the benefits of AI may not be fully realized, as users may resist or underutilize these systems, potentially leading to inefficiencies and a failure to achieve the intended improvements in parliamentary processes.

Guidelines for AI in parliaments

Generic risks and biases: Cognitive bias types

About this sub-guideline

This sub-guideline is part of the guideline Generic risks and biases. Refer to the main guideline for context and an overview. For a discussion of risks that relate more specifically to the unique work of parliaments, refer to the guideline Risks and challenges for parliaments.

This sub-guideline focuses on cognitive biases, which are systematic errors in judgements or decisions common to human beings owing to cognitive limitations, motivational factors and adaptations accumulated throughout life. Sometimes, actions that reveal cognitive biases are unconscious.

Automation bias

Automation bias occurs when conclusions drawn from algorithms are valued more highly than human analyses. For example, people will often blindly follow satellite navigation systems and arrive at the wrong place, or cross dangerous streets and put their life at risk.

Group attribution bias

Group attribution bias refers to the belief that everything that occurs for one individual is true for everyone. An example is when people stereotype professions with statements such as “all lawyers are manipulative” or “all artists are eccentric”.

Implicit bias

Implicit bias refers to the practice by which people unconsciously associate situations with their own mental model of representing that situation. For example, people often assume that a younger colleague cannot be experienced enough to be a good manager, or that an older employee is not able to learn new skills.

In-group favouritism

In-group favouritism occurs when someone acts with partiality towards the existing aspects of the group to which they belong. For example, people may systematically

recommend someone from their “group” for a job, while sports fans will always view their team as the best.

Out-group favouritism

Out-group favouritism refers to the favouring of groups outside the group to which a person belongs. For example, a manager who does not recognize the talent available in their own team will always turn to someone from another team for advice or support.

Affinity bias

Affinity bias happens when someone prefers individuals who are similar to them in terms of ideology, attitudes, appearance or religion. For example, a hiring manager might prefer a candidate who went to the same university as they did, overlooking other qualified applicants.

Social bias

Social bias occurs when many individuals in a society or community share the same bias. The simplest examples are religion and politics. Some people are so closed in a belief system that they are incapable of seeing both sides of an argument. They seek only information that supports their belief and negate anything that counters it, demonstrating their bias in their every action.

Rules and systems bias

Rules and systems bias refers to the fact that, when developers are used to particular rules embedded in systems, they try to reproduce the same rules to represent other situations. For example, developers sometimes choose solutions based on examples they readily remember. Controlled laboratory studies have identified the harmful effects of specific cognitive biases on several aspects of software development such as defect density, requirements specification, originality of design and feature design.

Requirement bias

Requirement bias refers to the assumption that all people or situations are capable of meeting, or meet, the same technical requirements (hardware and/or software). It is a subset of “rules and systems bias”.

Anchoring bias

Anchoring bias occurs when people rely too much on pre-existing information, or on the first information they find, when making decisions. For example, if someone sees a computer that costs $5,000, and then see a second one that costs $2,000, they are likely to view the second computer as cheap. This type of bias can impact procurement decisions.

Availability bias

Availability bias is a mental shortcut whereby people tend to ascribe excessive weight to what is readily “available” – i.e. what comes easily or quickly to mind –

when making judgements and decisions. For example, people remember vivid events like plane crashes over more common incidents such as car crashes, despite the latter being much more common. As a result, they often overestimate the likelihood that a plane will crash and might even choose to drive rather than fly, even though they are much more likely to be involved in a road traffic accident. This type of bias can occur when business staff are describing business rules to developers.

Confirmation bias

Confirmation bias refers to the fact that people tend to prefer information that confirms their existing beliefs. It affects how people design and conduct surveys, interviews or focus groups, and analyse competition. Essentially, people construct questions in a way that will produce the answers they want. For example, if someone types the question “Are dogs better than cats?” into an online search engine, articles that argue in favour of dogs will appear first. Conversely, the question “Are cats better than dogs?” will produce results in support of cats. This applies to any two variables: the search engine “assumes” that the person thinks variable A is better than variable B, showing them results that agree with their opinion first.

User interaction bias

User interaction bias occurs when a user imposes their own self-selected biases and behaviour when interacting with data, output or results. For example, when a system is trained using streaming data from a live group discussion, it instils the bias that exists in that group.

Groupthink

Groupthink refers to the fact that people in a group tend to make non-optimal decisions based on their desire to conform to the group, or for fear of dissenting. For example, when the leader of a group tells everyone that they need to ban all members of a particular ethnic group from joining them, the members of the group accept that decision without questioning it.

Funding bias

Funding bias occurs when biased results are reported in order to support or satisfy the organization funding a piece of research. For example, a study published in a scientific journal found that drinks containing high-fructose corn syrup did not increase liver fat or ectopic fat deposition in muscles. However, the “acknowledgements” section shows that one of the researchers received funding from a major soft-drinks company. The results may therefore have been skewed to paint the funding organization in a positive light.

Sunk cost fallacy

The sunk cost fallacy is a human tendency to continue with an endeavour or behaviour because resources such as money, time or effort have already been invested, regardless of whether the costs outweigh the benefits. For example, in AI, an organization that has already invested significant time and money in a particular AI application will pursue it to market rather than deciding to cancel the project, even in the face of significant technical or ethical debt.

Rashomon effect

The Rashomon effect is a term derived from the classic 1950 Japanese film *Rashomon*, which explores the concept of subjective reality and the nature of truth by presenting differing accounts of a single event from the perspectives of multiple characters. This bias occurs when there are differences in perspective, memory and recall, interpretation, and reporting on the same event from multiple witnesses. For example, people who attended a legislative committee meeting might have different perceptions regarding the debate and, therefore, provide a different summary of the event.

Streetlight effect

The streetlight effect refers to the fact that people tend to search only where it is easiest to look, such as when data scientists develop an AI algorithm using only a small data set (i.e. only the data they have access to) instead of considering obtaining more complete data from other organizations.

Ranking bias

Ranking bias is a form of anchoring bias. It refers to the fact that, in a list of search engine results, people believe that the highest-ranked results are the most relevant and important. They will still tend to click more on the top result than others, even if the results are ranked randomly.

Ecological fallacy

The ecological fallacy refers the drawing of conclusions about individuals based on group-level data. For example, if a specific neighbourhood has a high crime rate, people might assume that any resident living in that area is more likely to commit a crime.

Survivorship bias

Survivorship bias is when people focus on the items, observations or people who "survive" (i.e. make it past a selection process), while overlooking those who do not. For example, by assessing only "surviving" businesses and mutual funds, analysts record positively biased financial and investment information – omitting the many companies that failed despite having similar characteristics as the successful ones.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Generic risks and biases: Data bias types

About this sub-guideline

This sub-guideline is part of the guideline Generic risks and biases. Refer to the main guideline for context and an overview. For a discussion of risks that relate more specifically to the unique work of parliaments, refer to the guideline Risks and challenges for parliaments.

This sub-guideline focuses on data biases – a type of error in which certain elements of a data set are more heavily weighted or represented than others, painting an inaccurate picture of the population. A biased data set does not accurately represent a model's use case, resulting in skewed outcomes, low accuracy levels and analytical errors.

Selection bias

Selection bias occurs when selecting data.

In one example, an AI system for detecting Parkinson's disease was trained using a data set containing only 18.6% women. Consequently, the rate of accurate detection of symptoms was higher among male than female patients even though, in reality, the symptoms in question are more frequently manifested by female patients.

In another example, an AI system for detecting skin cancer was not able to detect the disease in people of African descent. Researchers observed that, as rates of skin cancer were increasing in Australia, the United States and Europe, the data set used to train the system consisted largely of people of European descent.

Sampling bias

Sampling bias is a form of selection bias in which data is not randomly selected, resulting in a sample that is not representative of the population. For example, if a poll for a national presidential election targets only middle-class voters, the sample will be biased because it will not be diverse enough to represent the entire electorate.

Coverage bias

Coverage bias is a form of sampling bias that occurs when a selected population does not match the intended population. For example, general national surveys conducted online may miss groups with limited internet access, such as the elderly and lower-income households.

Participation bias

Participation bias is a form of sampling bias that occurs when people from certain groups decide not to participate in the sample. It occurs when the sample consists of volunteers, which also creates a bias towards people who are willing and/or available to participate. The results will therefore only represent people who have strong opinions about the topic, omitting others.

Omitted variable bias

Omitted variable bias is a form of sampling bias that occurs when an important variable is omitted during data collection, compromising an expected result. For instance, when designing an algorithm that determines the price of second-hand cars, the developers include the following variables: make, number of seats, accident history, distance on the clock and engine size. However, they forget to include the car's age. The algorithm is likely to give biased estimates because two cars with exactly the same values for the other variables will probably have different prices according to their age.

Popularity bias

Popularity bias is a form of sampling bias that occurs when items that are more popular gain more exposure, while less popular items are underrepresented. For example, recommendation systems tend to suggest items that are generally popular rather than personalized picks. This happens because the algorithms are often trained to maximize engagement by recommending content that is liked by many users.

Data inaccuracy

Data inaccuracy is a result of failures in data entry. For example, with a system that registers temperature automatically, if there is a failure in the sensor, the data set will not be trustful for using temperature as a variable. Sometimes, systems are not rigid with data entry and accept data without standards or with errors.

Obsolete data

Obsolete data is data that is too old to reflect current trends. For example, a system designed to predict how long a public procurement exercise will take is trained on an excessively large data set, consisting mostly of procurement exercises that happened 10 years ago under different legislation. As a result, this system will likely produce inaccurate predictions.

Temporal bias

Temporal bias occurs when the training data is not representative of the current context in terms of time. For example, census data – which is only collected once every 10 years – is used for many predictions. However, if the last available census data was collected in 2021, i.e. in the middle of the COVID-19 pandemic, then algorithms that use this data may be biased in a number of ways.

Variable selection bias

Variable selection bias occurs when a chosen variable is not fit for purpose. For example, a national health agency looking to give an additional benefit to citizens selects, as the variable for allocation of the benefit, total health spending according to age. The algorithm selects people of European descent and those on higher incomes to receive the additional benefit. This biased outcome happened because people in this group spent more on their health. The chosen variable was the seed of the problem.

Confounding variable

A confounding variable, in research investigating a potential cause-and-effect relationship, is an unmeasured third variable that influences both the supposed cause and the supposed effect. For example, when researching the correlation between educational attainment and income, geographical location can be a confounding variable. This is because different regions may have varying economic opportunities, influencing income levels irrespective of education. Without controlling for location, it is impossible to determine whether education or location is driving income.

Simpson's paradox

Simpon's paradox is a phenomenon that occurs when subgroups are combined into one group. The process of aggregating data can cause the apparent direction and strength of the relationship between two variables to change. For example, a study shows that, within an organization, male applicants are more successful than women. However, comparing the rates within departments paints a different picture, with female applicants having a slight advantage over men in most departments.

Linguistic bias

Linguistic bias occurs when an AI algorithm favours certain linguistic styles, vocabularies or cultural references over others. This can result in output that is more relatable to certain language groups or cultures, while alienating others.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Generic risks and biases: Processing and validation bias types

About this sub-guideline

This sub-guideline is part of the guideline Generic risks and biases. Refer to the main guideline for context and an overview. For a discussion of risks that relate more specifically to the unique work of parliaments, refer to the guideline Risks and challenges for parliaments.

This sub-guideline focuses on processing and validation biases, which arise from systematic actions and can occur in the absence of prejudice, partiality or discriminatory intent. In AI systems, these biases are present in algorithmic processes used in the development of AI applications.

Aggregation bias

Aggregation bias arises when a model assumes a one-size-fits-all approach for different demographic groups that, in reality, may have different characteristics or behaviours.

Amplification bias

Amplification bias occurs when several AI systems, each with separate biases influenced by their training data and programming, interact and mutually reinforce each other's biases, leading to a more pronounced and persistent bias than what any single system might display.

For instance, a system trained on historical hiring data, in which male candidates have been predominantly selected, unintentionally favours male candidates during CV screening. Another AI system, tasked with performance evaluation, has been trained on data where female employees were often given lower scores owing to latent human biases. As these two systems interact, the hiring AI system may propose a larger number of male candidates, while the performance-evaluation AI system continues to judge female employees more harshly.

Deployment bias

Deployment bias – perhaps more of an operational failing than a bias – occurs when a system that works well in a test environment performs poorly when deployed in the real world owing to differences between the two environments.

Evaluation bias

Evaluation bias is a type of discrimination in which the methods used to evaluate an AI system's performance are biased, leading to incorrect assessments of how well the system is working.

Exclusion or sampling bias

Exclusion or sampling bias occurs when specific groups of user populations are excluded from testing and subsequent analyses.

Feedback loop bias

Feedback loop bias arises when the output of an AI system influences future inputs, potentially reinforcing and amplifying existing biases over time.

Model selection bias

Model selection bias is a technical term for confounding exploratory and hypothesis-testing statistical analyses. If data is used to select the best-fitting model from a set of candidates, that same data cannot then be used to test hypotheses about the value of the estimated parameters of the best-fitting model.

Optimization bias

Optimization bias occurs when the objective function of an AI system is defined in a way that leads to unintended consequences or unfair outcomes.

Overfitting or underfitting bias

Overfitting bias refers to a situation where a model is too complex and fits too closely to the training data, potentially incorporating noise or outliers that do not represent the true patterns in the data. Conversely, underfitting bias occurs when the model is too simple to capture the true patterns in the data, leading to poor performance and potentially biased results.

Proxy bias

Proxy bias occurs when variables used as proxies for protected attributes (such as race or gender) introduce bias into the model.

Temporal bias

Temporal bias occurs when training data becomes outdated and no longer represents current realities, leading to biased predictions. While this might be considered a data bias, it is also a processing/validation bias because it often occurs when systems fail to consider temporal aspects of the data validation process, or

when the process of updating and validating models fails to adequately account for changes in the underlying data distribution over time.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles

Audience

This guideline is intended for senior parliamentary managers and senior IT professionals in parliaments with responsibility for implementing AI and for its ongoing governance and management.

About this guideline

Parliamentary use of AI needs to be grounded in strong ethical practices. While this is true of AI use in any organization, it is particularly important for parliaments, which must ensure that they maintain public trust and confidence. Although the potential risks associated with AI use are generally agreed upon, managing and mitigating these risks requires an understanding of ethical principles.

This guideline and its sub-guidelines present a range of ethical principles related to AI. They discuss how AI can be implemented ethically across parliamentary processes and practice, at all levels of the institution. Ethical principles for AI are explored across eight areas:

- Privacy
- Transparency
- Accountability
- Fairness and non-discrimination
- Robustness and safety
- Human autonomy and oversight
- Societal and environmental well-being
- Intellectual property

Why ethical principles matter

In order to ensure that AI systems are trustworthy and used responsibly, parliaments should establish a code of ethics for the use of AI. This code should be applied during the use, development and deployment of AI systems, in order to manage or mitigate the risks of these technologies while maximizing their benefits.

The code of ethics should be explicit about what parliament expects from the operation of AI systems and from the people involved in their production and use. It should align with relevant national and international laws, regulations and standards. It should include recommendations, guidance and limitations for each ethical

principle, which should apply throughout the entire AI system life cycle – from planning to decommissioning.

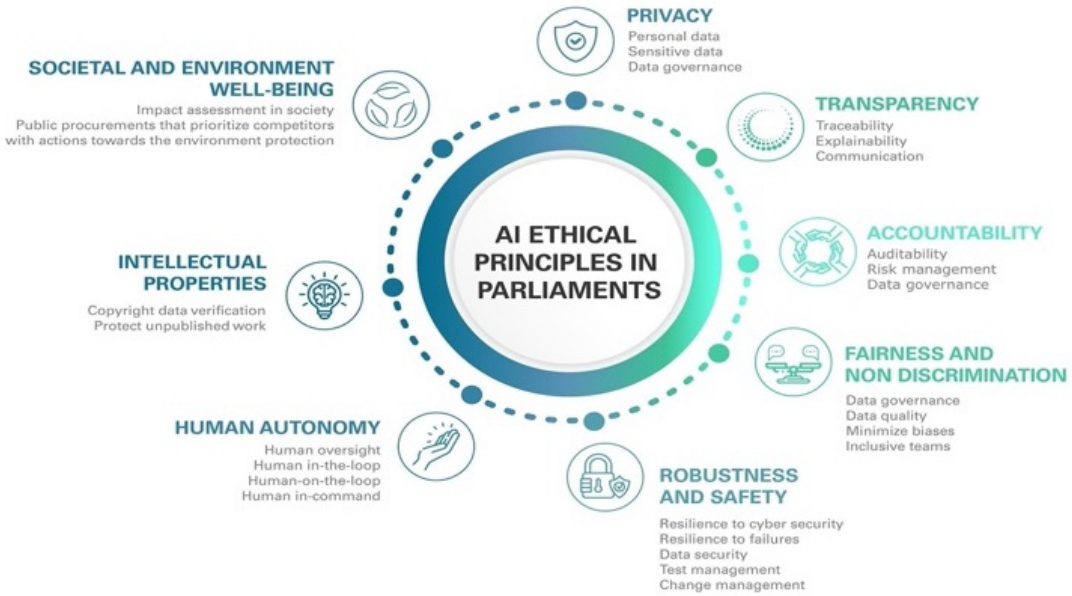
Developing ethical principles for parliaments

There are many resources available to parliaments to guide them in developing ethical principles for AI use. Some parliaments may already have a framework in place that can be adapted.

The following section presents a model that parliaments can adopt if they wish. In this model, ethical principles for parliamentary AI use are broken down into eight areas:

- **Privacy**: AI systems should respect and uphold privacy rights and data protection.
- **Transparency**: People should be able to understand when and how they are being impacted by AI, through transparency and responsible disclosure.
- **Accountability**: It should be possible to identify who is responsible for the different phases of the AI system life cycle.
- **Fairness and non-discrimination**: AI systems should be inclusive, accessible and not cause unfair discrimination against individuals, communities or groups.
- **Robustness and safety**: AI systems should reliably operate in accordance with their intended purpose.
- **Human autonomy and oversight**: AI systems should respect people's freedom to express opinions and make decisions.
- **Societal and environmental well-being**: AI systems should respect and promote societal well-being and environmental good.
- **Intellectual property**: AI systems should respect intellectual property rights.

Figure 1: Ethical principles for parliaments



From minimizing biases and ensuring robust oversight to maintaining clear communication and protecting personal data, these principles work together to create a comprehensive framework for ethical AI use. By adhering to these principles, parliaments can harness the benefits of AI while mitigating risks, fostering public trust and upholding their democratic responsibilities.

Each of these areas is explored in turn in the remainder of this guideline, and in the associated sub-guidelines, which describe the specific challenges and considerations for parliaments, and offer practical guidance, actionable strategies and recommendations.

Privacy

This sub-guideline explores the principle of privacy in AI governance for parliaments, with a focus on personal data protection. It outlines specific privacy concerns in various parliamentary work processes, including legislative, administrative and citizen interaction contexts. It emphasizes the importance of justifying and limiting the use of personal data in AI systems, and provides guidance on handling sensitive information. Special attention is given to the challenges posed by generative AI in processing personal and sensitive data.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that respect privacy and protect personal data.

For further guidance on the principle of privacy, refer to the sub-guideline Ethical principles: Privacy.

Transparency

This sub-guideline explores the principle of transparency in AI governance for parliaments. It defines transparency as the communication of appropriate information about AI systems in an understandable and accessible format. The sub-guideline addresses three key aspects of transparency: traceability, explainability and communication.

Highlighting the importance of documenting the entire life cycle of AI systems, from planning to decommissioning, it provides practical recommendations for implementing transparency. These include risk assessment documentation, standardized methods for explaining AI decisions, and clear communication about AI system capabilities and limitations. The sub-guideline also offers specific guidance on ensuring transparency in generative AI applications, acknowledging the unique challenges they present.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are transparent, accountable and aligned with democratic values.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Transparency.

Accountability

This sub-guideline explores the principle of accountability in AI governance for parliaments. It emphasizes that while AI systems themselves are not responsible for their actions, clear accountability structures are essential.

The sub-guideline discusses the importance of auditability and risk management throughout the AI system life cycle. It provides practical recommendations for implementing accountability, including stakeholder identification and risk assessment processes, and for preparing for both internal and external audits.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are accountable and aligned with democratic values.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Accountability.

Fairness and non-discrimination

This sub-guideline explores the principle of fairness and non-discrimination in AI governance for parliaments, including minimizing biases in legislative processes and citizen interactions. It emphasizes the importance of trust and provides specific recommendations for dealing with potential biases.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are fair, non-discriminatory and free from biases.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Fairness and non-discrimination.

Robustness and safety

This sub-guideline explores the principle of robustness and safety in AI governance for parliaments, emphasizing that, in order to be trustworthy, AI systems should be robust to adversity and to changes within the environment for which they were designed.

The sub-guideline presents the principle of robustness and safety through two lenses: resilience to failures that could cause damage to people, organizations or the environment or that could prevent traceability, and resilience to cyberattacks.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are robust and safe.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Robustness and safety.

Human autonomy and oversight

This sub-guideline explores the principle of human autonomy and oversight in AI governance for parliaments. It refers to the way in which AI systems interact with

humans, as well as the way in which information is stored, transmitted and secured. It stresses that parliaments, as enablers of a democratic, flourishing and equitable society, must support the user's agency and uphold fundamental rights and that, in an AI context, this requires human oversight.

In this sub-guideline, special attention is given to the challenges posed by generative AI, emphasizing the need for robust feedback channels and frequent human checks.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that promote human autonomy and allow for human oversight.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Human autonomy and oversight.

Intellectual property

This sub-guideline explores the principle of intellectual property in AI governance for parliaments. It emphasizes that everyone involved in an AI system's life cycle, including users, must respect intellectual property in order to protect the investment of rights-holders in original content. It covers copyrights, accessory rights, and contractual restrictions on accessing and using content.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that respect intellectual property rights.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Intellectual property.

Societal and environmental well-being

This sub-guideline explores the principle of societal and environmental well-being in AI governance for parliaments. It emphasizes that, owing to the ubiquitous nature of AI in society, this technology should be used for people's well-being. It further stresses that applications of AI should not negatively affect people's physical and mental well-being or harm the environment or society at large.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that protect and promote societal and environmental well-being.

For further guidance on the principle of transparency, refer to the sub-guideline Ethical principles: Societal and environmental well-being.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Privacy

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of privacy in AI governance for parliaments, with a focus on personal data protection. It outlines specific privacy concerns in various parliamentary work processes, including legislative, administrative and citizen interaction contexts. It emphasizes the importance of justifying and limiting the use of personal data in AI systems, and provides guidance on handling sensitive information. Special attention is given to the challenges posed by generative AI in processing personal and sensitive data.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that respect privacy and protect personal data.

Why privacy matters

In the context of digital transformation, it is important – and often a legal requirement – to protect personal data. AI systems are no exception to this rule, with parliaments needing to comply with both legislation and internal standards on this subject.

Before embarking on an AI project, parliaments will therefore need to delve deeper into the problem to be solved with AI in order to identify whether any type of personal data will be collected, processed and potentially shared, clearly recording the classification of the data and its precise meaning.

The following sections address specific privacy concerns raised by the diverse work processes of parliaments.

Use of personal data

It is important to exercise caution and restraint when using personal data, and only to do so where absolutely necessary. Parliaments should adhere to the following principles:

- Where personal data is deemed essential to an AI system's functionality, a clear and compelling justification for its use must be made. This justification should be subject to rigorous scrutiny and approval by a data protection

officer (or equivalent person) and by key decision makers within the AI governance framework.
- If approval for the use of personal data is given, strict practices must be put in place to safeguard privacy and to prevent misuse. Such practices must protect individuals from exposure, even indirectly, especially when dealing with biometric data or when combining information from multiple sources.
- AI systems must not profile individuals according to their behaviour or use personal data in ways that could lead to discrimination, the manipulation of opinions, or any form of harm, whether psychological, physical or financial.
- Explicit authorization should be required for the use of sensitive data, adding an extra layer of protection and accountability.
- Special conditions may be required for the use of personal data for research purposes or to support bills going through parliament, especially if parliament already has internal regulations regarding the use of personal data.

Administrative processes

Where parliament is adopting or developing AI systems, it should identify, understand and document what data is being used – both internal data, and externally sourced or hosted data – and identify who the owner of that data is.

Citizens' data

When interacting with citizens, parliaments must take special care to manage and protect the personal data they collect, such as through an online digital service or a manual data-collection process. They must also carefully consider what data is stored in a system that is exposed to AI, and ensure that only essential data is retained. More generally, when designing an AI system, parliaments need to understand the parameters of data privacy, knowing what is admissible for release into the public domain, what must be anonymized, and what is protected.

Sensitive data and generative AI

Parliaments must exercise extreme caution and appropriate scrutiny when feeding personal and sensitive data into generative AI systems, as these systems will process and use any data given to them. The institution should have in place mechanisms to protect its personal and sensitive data from inadvertent or inappropriate access by such tools. This is especially important if this data is processed externally, as is the case with most generative AI systems.

Where a parliament does authorize personal data for use by generative AI systems, it should actively implement processes to anonymize this data, as well as adopting other mechanisms, established by internal rules, before submitting any personal data to such tools. This practice minimizes the risk of personal data breaches and misuse.

Practising privacy

In order to ensure that AI systems respect and protect privacy, parliaments should adopt a comprehensive approach. The components of this approach are detailed below:

- Conduct a thorough assessment of AI systems to identify any use of personal data, clearly documenting data classification and purpose.
- Implement strict data protection practices, including obtaining approval from the data protection officer for any use of personal data in AI systems.
- Establish clear protocols for managing citizens' data in AI-driven interactions, ensuring that only essential data is collected and stored.
- Develop and enforce stringent safeguards for handling sensitive data, particularly when using generative AI tools.
- Create a comprehensive data ownership and management system, documenting both internal and external data sources used in AI processes.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Transparency

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of transparency in AI governance for parliaments. It defines transparency as the communication of appropriate information about AI systems in an understandable and accessible format. The sub-guideline addresses three key aspects of transparency: traceability, explainability and communication.

Highlighting the importance of documenting the entire life cycle of AI systems, from planning to decommissioning, it provides practical recommendations for implementing transparency. These include risk assessment documentation, standardized methods for explaining AI decisions, and clear communication about AI system capabilities and limitations. The sub-guideline also offers specific guidance on ensuring transparency in generative AI applications, acknowledging the unique challenges they present.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are transparent, accountable and aligned with democratic values.

Why transparency matters

Transparency involves communicating appropriate information about AI systems to the right people and in a free, understandable and easily accessible format.

Transparency – throughout the entire life cycle of an AI system – encompasses three key aspects: traceability, explainability and communication. These are discussed below.

Traceability

Traceability implies the ability to follow and monitor the entire life cycle of an AI system, from the definition of its purpose, through to planning, development, use and ultimate decommissioning.

Architects, developers, decision makers and even users involved in the development and evolution of AI systems are advised to use a combination of tools and documentation to support traceability.

Explainability

Explainability is the ability for humans to understand and trust each decision, recommendation or prediction made by an AI system.

As complexity increases in AI systems, explainability declines. Consequently, initially simple AI systems become less explainable as new layers of functionality are added over time.

Since different AI system stakeholders require different types of explanations, parliaments must generate documentation aimed at decision makers and those responsible for AI governance, in addition to the documents produced by the development team.

Communication

Communication is important for transparency: humans must always know that they are interacting with an AI system. As such, any AI system that interacts with humans must identify itself unambiguously. It must be explained to users and practitioners, in a clear and accessible manner, how the system functions and what its limitations are.

Practising transparency in AI systems

In order to ensure that AI systems are transparent, parliaments should adopt a comprehensive, life cycle-wide approach. The components of this approach are detailed below:

- **Risk assessment**: Produce a comprehensive risk assessment to guide project authorization, development and maintenance. This assessment should consider all stakeholders and inform decisions from initiation to potential decommissioning.
- **Standardization**: As part of the AI systems development process, adopt a standardized transparency method, such as Explainable AI (XAI), to document key aspects including problem definition, data selection criteria, personal data usage, technical specifications, user feedback and oversight results. Capture the rationale behind all significant decisions.
- **Reporting**: Maintain transparency through regular behaviour reports and continuous data storage for auditing. Clearly communicate system expectations, limitations and potential abnormalities to all relevant parties.

- **Communication**: Ensure that AI applications interacting with humans disclose their artificial nature. Inform business managers about AI usage in their areas of responsibility.
- **Documentation**: Tailor transparency documentation to the intended audience, whether internal or external. For outsourced AI systems, clearly communicate and enforce transparency requirements with external providers.

Practising transparency in generative AI systems

Parliaments using generative AI must prioritize transparency and responsibility, recognizing that the learning process and data used by AI systems may not be transparent:

- Label AI-assisted documents, specifying the tool and version used.
- Establish clear guidelines for permissible AI use in document creation.
- Document AI processes from design to deployment, including mechanisms for ensuring trust in AI systems.
- Prioritize commercial AI tools aligned with human rights frameworks.
- Justify and document any use of personal or third-party data.
- Clearly communicate when AI outputs are probabilistic rather than factual.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Accountability

About this sub-guideline

This sub-guideline on accountability is part of the guideline on the ethical principles for the use of AI in parliaments.

It explores the principle of accountability in AI governance for parliamentary settings. The sub-guideline discusses the importance of auditability and risk management throughout the AI system lifecycle. It provides practical recommendations for implementing accountability, including stakeholder identification, risk assessment processes, and preparing for both internal and external audits.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are accountable, fair, and aligned with democratic values.

Why accountability is important

The accountability principle in AI governance focuses on establishing clear structures, processes, and tools to evaluate and hold AI systems accountable, as the systems themselves cannot be responsible for their actions. Parliaments should ensure clear accountability for all decisions and actions throughout an AI system's lifecycle, from planning to decommissioning.

Effective accountability relies on two key elements:

1. Auditability: The ability to track the entire process of an AI system's lifecycle, including planning, development, use, and maintenance. This depends heavily on transparency and may involve both internal and external audits.
2. Risk Management: The identification, evaluation, documentation, and minimisation of risks associated with AI systems. This proactive approach helps identify potential vulnerabilities and assigns responsibility for risk mitigation.

These practices are particularly important for parliaments that undergo frequent audits. By implementing robust accountability measures, parliaments can ensure

their AI systems remain trustworthy and aligned with their democratic responsibilities.

Practices towards accountability

To ensure accountability in AI systems in parliament, it's crucial to adopt a comprehensive approach throughout the system's lifecycle. Begin by identifying all the stakeholders affected by the AI system, whether directly or indirectly. This holistic view helps anticipate potential impacts and concerns.

Next, implement a robust risk management process. This should encompass identifying, evaluating, documenting, minimising and continuously monitoring risks associated with AI systems. Such a process allows for the proactive management of potential issues before they escalate.

Establish rigorous internal auditing processes for AI systems. These regular checks help maintain system integrity and provide ongoing assurance of compliance with ethical standards and operational requirements.

It's equally important to prepare staff for external audits. Provide thorough training to equip team members with the knowledge and skills needed to engage confidently with third-party auditors, ensuring transparency and cooperation.

Finally, conduct a thorough assessment to identify which AI systems require trustworthy certification. Once identified, develop a clear, actionable plan to achieve this certification. This step not only enhances the system's credibility but demonstrates a commitment to maintaining high standards of AI governance.
By implementing these measures, parliaments can create a culture of accountability around their AI systems, fostering trust and ensuring responsible deployment of this powerful technology.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Fairness and non-discrimination

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of fairness and non-discrimination in AI governance for parliaments, including minimizing biases in legislative processes and citizen interactions. It emphasizes the importance of trust and provides specific recommendations for dealing with potential biases.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are fair, non-discriminatory and free from biases.

Why fairness and non-discrimination matter

Fairness can be defined at the individual level (such ensuring similar individuals are treated consistently) or at the group level. In the latter case, this involves grouping people into categories and ensuring that these groups are treated equitably.

Fairness, in the context of AI, is the ability of AI systems to not discriminate or reinforce biases against any individual or group. This principle is based on impartiality and inclusion. Fair, non-discriminatory decisions therefore presuppose bias-free data and algorithms.

Practising fairness and non-discrimination

In order to ensure that AI systems are fair and non-discriminatory, parliaments should adopt a comprehensive approach to data management and bias mitigation. The components of this approach are detailed below:

- **Data quality management**: Establish robust processes to manage data quality, particularly for data sets likely to be used in AI systems. Implement practices to ensure that there are no biases in the data and in the models that

will be used to train the algorithms. Such practices should consider not only data biases and processing biases, but also cognitive biases (for further guidance on this topic, refer to the guideline Generic risks and biases and its associated sub-guidelines).
- **Staff training**: Provide staff with training in data ethics, focusing on identifying and minimizing biases throughout the AI development process.
- **Data governance**: Implement a data governance process, with a clear delineation of responsibilities between data owners and data stewards.
- **Collaboration**: Have IT and business units work closely together. Such collaboration is vital for predicting, minimizing and monitoring biases throughout the AI system life cycle.
- **Data ethics committee or team**: Establish a data ethics committee or a multi-skilled team capable of analysing potential biases and communicating them to both managers and IT teams for each AI project.
- **Diversity and inclusivity**: Prioritize diversity and inclusivity when forming project teams and data ethics committees. By bringing together individuals of different ages, genders, ethnicities and skill sets, parliaments can ensure that a broad range of perspectives are heard, reducing the risk of that potential biases could be overlooked and enhancing the overall fairness of AI systems.

Minimizing biases in parliamentary processes

When planning and developing AI systems for use in legislative processes, parliaments should:

- Ensure that the data does not contain biases regarding political-party ideology and previous value judgements
- Be aware of possible historical biases in data relating to committee meetings and plenary sessions
- Establish partnerships with public organizations from which they regularly source external data for AI-powered bill-drafting systems, in order to maintain data quality
- Be aware of biases in text translation and speech-to-text transcription
- Confirm whether the information produced by generative AI systems is free from biases before considering using them

When planning and developing AI systems for use in government oversight processes, parliaments should:

- Identify data quality problems in government data and alert the government agency in charge of the data
- Establish partnerships with government agencies in charge of the data in order to improve data quality and minimize biases

When planning and developing AI systems for use in citizen interaction processes, parliaments should:

- Identify biases coming from citizens
- Avoid internalizing biases presented by citizens
- Avoid exposing any biases when interacting with citizens

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Robustness and safety

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of robustness and safety in AI governance for parliaments, emphasizing that, in order to be trustworthy, AI systems should be robust to adversity and to changes within the environment for which they were designed.

The sub-guideline presents the principle of robustness and safety through two lenses: resilience to failures that could cause damage to people, organizations or the environment or that could prevent traceability, and resilience to cyberattacks.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that are robust and safe.

Resilience to cyberattacks

Cyberattacks against AI systems exploit both algorithmic opacity and the strong dependence of algorithms on data. Such attacks may be difficult to detect in a timely manner, requiring AI systems security management practices that encompass advanced prevention techniques and focus on restoring the system and the entire environment to normal operating conditions.

Resilience to failure

Failures can occur in AI systems when variables take on unknown or false values that the developer did not consider and did not programmatically act to prevent. While AI systems are generally expected to be robust, if such failures do occur, there must be a mechanism to restore the system to its normal state in a timely and responsible manner, with minimal loss of data or impact on parliament.

Practising robustness and safety

In order to ensure that AI systems are robust and safe, parliaments should adopt a comprehensive, dynamic risk management approach that adapts to the ever-changing environment in which these systems operate. The components of this approach are detailed below:

- **Comprehensive testing**: Identify and mitigate cyber threats specifically targeting AI systems, while not neglecting other potential vulnerabilities.
- **Security practices**: Tailor security practices to address the unique challenges posed by AI systems and the threats they face, including through close and rapid communication between data teams and information security experts. When developing AI systems, cybersecurity should be a primary consideration, integrated from the outset, rather than added as an afterthought.
- **Training**: Invest in continuous training for developers and information security staff. These staff should be well-versed in techniques to prevent cyberattacks on AI systems and equipped with disaster recovery strategies specific to these technologies.
- **Internal collaboration**: Ensure that internal business units responsible for AI systems work closely with IT departments to establish clear parameters for monitoring system behaviour and defining thresholds for alerts regarding suspicious activity.
- **External partnerships**: Forge partnerships with other public institutions. These alliances facilitate swift and effective communication about emerging threats and new attack categories. They also provide a platform for sharing experiences – both successes and failures – in implementing various security techniques and technologies.

By adopting this holistic approach, parliaments can create a resilient framework for AI systems that can withstand threats, adapt to changes, and continue to serve their intended purpose effectively and safely.

Maintaining safety when using generative AI

When implementing generative AI in parliamentary contexts, safety considerations are paramount:

- Maintain strict control over data access, ensuring that AI systems and tools only interact with data specifically authorized for their intended purpose. This approach safeguards sensitive information and maintains the integrity of parliamentary processes.
- Where data transfer to external cloud services raises security concerns or presents other risks, explore alternative solutions. One viable option is to employ open-source generative AI models that can run locally on a parliament's own systems. This strategy provides the benefits of generative AI while offering full control over security, data management and integrity.

By adopting these measures, parliaments can harness the power of generative AI while upholding the highest standards of data protection and operational safety.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Human autonomy and oversight

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of human autonomy and oversight in AI governance for parliaments. It refers to the way in which AI systems interact with humans, as well as the way in which information is stored, transmitted and secured. It stresses that parliaments, as enablers of a democratic, flourishing and equitable society, must support the user's agency and uphold fundamental rights and that, in an AI context, this requires human oversight.

In this sub-guideline, special attention is given to the challenges posed by generative AI, emphasizing the need for robust feedback channels and frequent human checks.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that promote human autonomy and allow for human oversight.

Human autonomy and protection of citizens' rights

Humans should be free to express their opinions and make decisions about their lives without interference, coercion or manipulation. In order to ensure that AI systems do not negatively affect citizens' rights, it is important for parliaments to understand how AI systems interact with humans and how information is stored, transmitted and secured.

Maintaining human autonomy is particularly challenging in certain areas, especially web searches, content curation, content moderation, browsing activities, and email and text communications stores in the cloud. When parliaments rely on AI systems to execute core public functions, they should ensure that the design and operation of these systems complies with international human rights standards, as part of their duty to promote freedom of expression.

Moreover, in order to preserve human agency, direct interaction between human end users and an AI system should be established in a way that avoids simulating social relationships or stimulating potentially negative or addictive behaviour.

Human oversight

In the day-to-day running of organizations, human oversight is exercised through human supervision of an AI system's outputs. Users and managers responsible for AI systems analyse this output to ascertain whether undesirable behaviours have occurred, whether the rules established at the development stage need to be modified, or whether there are any data biases that went unnoticed during the development of the system.

The nature of human oversight depends on the type of AI application. Supervision can occur during development, on an ongoing basis, or once the system is in production, in order to gather feedback on the system's output.

There are three types of human supervision that can be applied to AI systems:

- **Human-in-the-loop (HITL)**: Under this model, a human mediates all decisions made by the AI system. While this approach offers the highest level of human control, it is not always desirable or feasible, particularly for systems designed for rapid decision-making or high-volume data processing.
- **Human-on-the-loop (HOTL)**: This approach allows for human intervention during the project development phase. Once the system is operational, the human's role shifts to monitoring the system's operation and decisions. This approach balances automation with human oversight, allowing for intervention when necessary.
- **Human-in-command (HIC)**: This is the most comprehensive form of oversight, extending beyond the AI system's immediate functioning to consider broader economic, social, legal and ethical impacts. Under this model, oversight can even extend to society at large, with public feedback gathered on the AI system's behaviour providing a broader perspective on its effects and implications.

The distinction between these three approaches – HITL, HOTL and HIC – lies primarily in the level of autonomy granted to the AI system and the extent of human oversight. These are summarized in Table 1 below:

Table 1: Models of Human-AI interaction

	HUMAN-IN-THE-LOOP (HITL)	HUMAN-ON-THE-LOOP (HOTL)	HUMAN-IN-COMMAND (HIC)
Level of Human Involvement	Continuous; collaborative; active participation in decision-making	Supervisory; intervene only when necessary	Ultimate decision-making authority
AI Autonomy	Reliant on human handshake; varies; typically acts autonomously until human review is needed	Operates autonomously with human oversight	Can act autonomously but will never decide autonomously
Efficiency	Lower, due to the need for constant human input	Higher than HITL, balanced with oversight	Varies; prioritizes control over efficiency
Control & Safety	High control; allows for nuanced decisions	Balanced control; efficient for routine tasks	Maximum control

By implementing these oversight approaches, parliaments can harness the benefits of AI while maintaining essential human control and accountability, thus upholding democratic principles and public trust.

Practising human autonomy and oversight

In order to safeguard human autonomy and maintain proper oversight of AI systems, parliaments should adopt a comprehensive approach. The components of this approach are detailed below:

- **Risk assessment**: Conduct a thorough risk assessment of each AI system, paying particular attention to those that interact directly with human end users. For these systems, identify any potential for confusion about who or what is engaging in the interaction.
- **Rules**: Establish clear rules for AI-human interactions in order to prevent any manipulation or the formation of inappropriate social relationships. The type of oversight required for each AI system should be determined according to its specific risk profile.
- **Standards and testing**: Develop a set of clear, measurable criteria for acceptable and unacceptable AI behaviours, and draw up an extensive testing plan to explore the full range of system behaviours. Parliament's AI policy should designate a specific position or organizational body with the authority to withdraw an AI system from operation if it cannot meet these behavioural standards.
- **Training**: Provide thorough training on the assessment process to both technical staff and managers, including on the use of any specific tools or functionalities built into the AI system itself.
- **Reporting**: Produce regular oversight reports, appropriately tailored for both technical staff and managers.
- **Review**: Establish a timely and efficient process for reviewing these human oversight reports to ensure that any issues or concerns are addressed promptly, thus maintaining the integrity and trustworthiness of the AI systems in use.

By implementing these practices, parliaments can leverage the capabilities of AI systems while ensuring that these remain under appropriate human control and respect human autonomy.

Human oversight of generative AI

Parliaments should establish robust oversight mechanisms for generative AI:

- Create a digital channel for user feedback on AI outputs.
- Invest in staff training for effective AI oversight.
- Conduct more frequent human checks on AI-generated content, given the rapid advancements in this technology.
- Carefully select generative AI tools and ensure that all users understand their specific limitations.

These measures allow parliaments to leverage generative AI while maintaining essential human control and ethical standards.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Societal and environmental well-being

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of societal and environmental well-being in AI governance for parliaments. It emphasizes that, owing to the ubiquitous nature of AI in society, this technology should be used for people's well-being. It further stresses that applications of AI should not negatively affect people's physical and mental well-being, or harm the environment or society at large.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that protect and promote societal and environmental well-being.

Practising societal and environmental well-being

In order to ensure that AI systems promote and support societal and environmental well-being, parliaments should adopt a comprehensive approach. The components of this approach are detailed below:

- **Stakeholder identification**: Identify potentially impacted stakeholders (individuals, organizations and the environment) during AI system planning.
- **Procurement**: Prioritize environmentally responsible suppliers in AI-related public procurement exercises.
- **Impact assessment**: Assess the impacts of an AI system on all stakeholders throughout its life cycle.
- **Feedback**: Implement user feedback mechanisms to gauge an AI system's impact on work processes.
- **Citizen interaction**: Ensure unbiased and accurate AI interactions with citizens.
- **Misinformation**: Check AI-generated communication outputs for misinformation.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Ethical principles: Intellectual property

About this sub-guideline

This sub-guideline is part of the guideline Ethical principles. Refer to the main guideline for context and an overview.

This sub-guideline explores the principle of intellectual property in AI governance for parliaments. It emphasizes that everyone involved in an AI system's life cycle, including users, must respect intellectual property in order to protect the investment of rights-holders in original content. It covers copyrights, accessory rights, and contractual restrictions on accessing and using content.

Overall, this sub-guideline provides a framework for parliaments to develop and maintain AI systems that respect intellectual property rights.

Protecting intellectual property in AI systems

In order to protect intellectual property in AI systems, parliaments should adopt a comprehensive approach. The components of this approach are detailed below:

- Before starting to develop an AI system, check for copyrighted data and any contractual conditions.
- When using generative AI, check whether the presented data sources generate copyrighted content.
- Standardize the types of documents for which generative AI cannot be used to generate content.
- Expressly inform readers if a document's content has been written using generative AI.
- Protect parliament's unpublished or sensitive work by avoiding uploading it into an online AI system unless there are assurances that the data will not be reused.
- Train internal staff on the technical and ethical implications of intellectual property rights.
- Create a specific peer-review process for researchers who are using generative AI, in order to both maintain high standards of quality and protect against breaches of intellectual property rights.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Introducing AI applications

Audience

This high-level guideline is intended for senior parliamentary managers, as well as for parliamentary staff and MPs who are interested in gaining a broad understanding of where AI can impact upon the work of parliaments.

About this guideline

This guideline outlines considerations and recommendations for parliaments facing new challenges in procuring, implementing and managing AI systems at a time when this technology is becoming increasingly prevalent in standard software packages. It addresses how parliaments can approach AI-enhanced off-the-shelf products while adhering to the ethical principles and risk management strategies outlined in these Guidelines.

Procurement considerations

Increasingly, many off-the-shelf software packages are being augmented with AI functionalities. This AI integration is often opaque, and it is not immediately apparent to the user what impact AI is having, or how it works behind the scenes:

- Microsoft 365 and Microsoft Edge products are starting to embed Microsoft Co-pilot AI support.
- Google Docs already has new AI features such as “Help me write”, “Smart compose”, “Summarization” and “Voice typing”.
- Adobe Acrobat has integrated several AI functionalities, including AI Assistant for generating summaries and creating multi-document insights.
- Photo and video editing software often contains AI augmentation that makes it easy to render manipulated images.

When procuring AI-enhanced products, parliaments must therefore exercise due diligence to ensure alignment with their ethical standards and operational needs.

Checklist:

- Conduct thorough vendor assessments, focusing on AI ethics and data practices.

- Evaluate AI features against parliamentary needs and ethical guidelines.
- Ensure contracts include robust clauses on data protection and AI accountability.

Implementation strategy

A measured approach to implementation allows parliaments to assess the impact of AI functionality and its alignment with existing processes. Starting with a pilot phase provides an opportunity to develop clear protocols for AI feature usage and establish monitoring mechanisms.

Checklist:

- Begin with a pilot phase to evaluate AI impact and alignment.
- Develop protocols for enabling or restricting AI features based on task sensitivity.
- Establish monitoring mechanisms for AI-driven decisions or suggestions.

User training and awareness

Comprehensive training is crucial to ensure staff can leverage AI features effectively while understanding their limitations and potential risks. This is particularly important given the ethical considerations and potential biases inherent in AI systems.

Checklist:

- Develop training programmes highlighting AI benefits and risks (including data literacy and AI literacy).
- Create user guidelines for the appropriate usage of AI functionalities.
- Educate staff on recognizing and critically evaluating AI-generated content.

Ongoing management and ethical considerations

As AI capabilities evolve, parliaments must regularly review and update their policies and practices. This includes maintaining open communication with vendors, conducting periodic audits, and ensuring ongoing alignment with ethical principles such as fairness, transparency and human oversight.

Checklist:

- Regularly review and update AI usage policies.
- Maintain open communication channels with vendors.
- Conduct periodic audits of AI feature usage and impact.
- Assess and mitigate potential biases in AI-enhanced features.
- Ensure transparency and maintain human oversight in AI use.

Data governance and performance evaluation

Strict data governance is essential when using AI-enhanced products, particularly given the sensitive nature of parliamentary work. Regular performance evaluations help to ensure that AI features continue to meet parliamentary needs and ethical standards.

Checklist:

- Implement strict data access controls for AI-enhanced features.
- Ensure compliance with data protection regulations and parliamentary policies.
- Regularly assess the effectiveness and appropriateness of AI features.
- Gather user feedback to inform future procurement and implementation decisions.

Conclusion

By following these recommendations and checklists, parliaments can harness the benefits of AI-enhanced off-the-shelf products while mitigating risks and upholding ethical standards. This approach aligns with the broader AI governance framework outlined in the Guidelines, ensuring a consistent and responsible approach to AI adoption across all areas of parliamentary work.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Training for data literacy and AI literacy

Audience

This guideline is intended for a diverse audience within parliaments, including staff and MPs without prior AI expertise, senior staff responsible for AI implementation, decision makers at various levels (particularly those responsible for training and staff development budgets), users of parliamentary technologies, and technical staff involved in AI production and procurement.

About this guideline

This guideline focuses on the importance of awareness-raising and training for both data literacy and AI literacy in parliaments. It outlines the need for a comprehensive training programme tailored to different roles within parliament, including decision makers, users, technical staff and MPs, providing detailed recommendations for training content in both areas.

The AI literacy sub-guideline includes the fundamentals of AI, machine learning basics, ethical implications and practical applications in governance. The data literacy sub-guideline, meanwhile, encompasses topics such as data management principles, data collection, cleaning and analysis, and ethical considerations.

This guideline stresses the importance of understanding AI's potential benefits and risks in a parliamentary context, as well as the ethical considerations involved. It also highlights the need for a multidisciplinary approach and the development of a data-driven culture within parliaments.

Overall, this guideline provides guidance for parliaments on building capacity and preparing their staff for the effective and responsible use of AI technologies.

The need for training

The transformative potential of AI in parliaments necessitates comprehensive staff preparation and training. All parliamentary staff – from technical implementers to end users – will benefit greatly from understanding AI in general and, more specifically, its impact on their roles. Knowledge of the related ethical considerations and complexities is also important.

By strategically building staff skills and capabilities, parliaments can maximize the benefits of AI while effectively managing its risks, avoiding both overestimation of AI's impact and underestimation of its challenges.

Given that AI systems rely on data, parliamentary training programmes should cover both AI literacy and data literacy – since good-quality, well-managed and understood data is at the heart of successful AI implementations. It is important for parliaments to start developing these programmes before working with AI.

What is data literacy?

Data literacy is the ability to read, understand, create and communicate data as information. It involves understanding how to effectively collect, analyse, interpret and present data in meaningful ways. Data literacy includes knowing where data comes from, and grasping basic statistical concepts and data presentation and visualization techniques. Data literacy is vital for critically evaluating data-driven arguments and conclusions.

Those requiring a more advanced level of data literacy will need to understand the following topics:

- Defining the strategic questions to be answered by data
- Identifying and analysing the dependencies between questions and data
- Designing data architectures
- Implementing data governance and data management
- Analysing and processing data, including practices for data ethics
- Building digital products based on data analysis, including dashboards and data visualizations

What is AI literacy?

AI literacy is an understanding of the basic principles, capabilities and limitations of AI – something that is crucial for informed decision-making about AI adoption and oversight in parliaments. It involves the ability to recognize AI applications, grasp fundamental concepts like machine learning and data analysis, and critically evaluate AI's potential impacts. An AI-literate workforce can better leverage AI to enhance parliamentary functions while identifying and mitigating risks, ensuring responsible AI use that aligns with democratic principles.

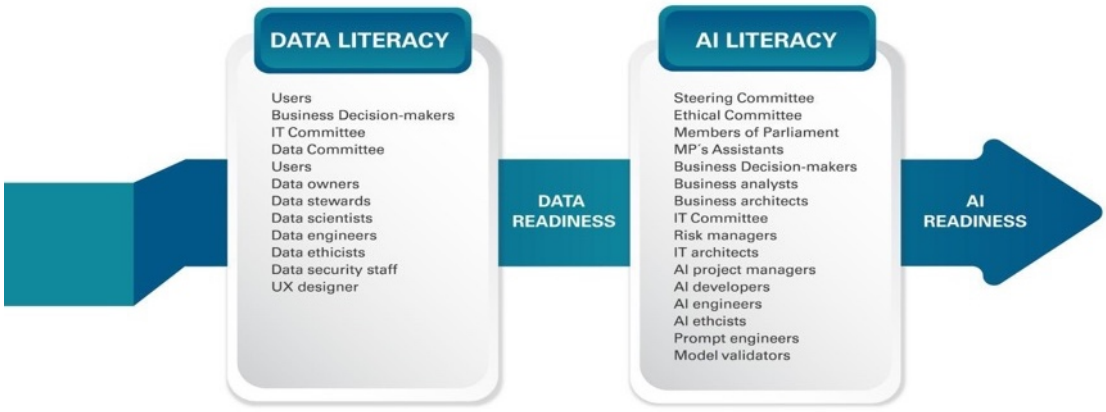
A more advanced level of AI literacy encompasses a deep understanding of the following:

- The fundamental principles of AI
- The risks of AI systems, as well as the resulting business value and outcomes
- Technology and applications
- Analytical and algorithmic methods
- Data and most dimensions of data literacy
- Ethical considerations regarding AI

Relationship between data literacy and AI literacy

Data readiness is a prerequisite for AI readiness in parliaments and requires multidisciplinary working teams across the product portfolio and throughout the system's life cycle. It is important for all AI initiatives to consider data governance and management processes and, as Figure 1 below shows, for data literacy to be a prerequisite for AI literacy training.

Figure 1: Possible roles to be considered in AI literacy and data literacy training programmes



The figure above emphasizes the importance of starting a data literacy programme before working with AI, and why it is important that parliaments have a plan to ensure good levels of AI literacy.

In addition, the emergence of generative AI takes the technology to the end user and gives them the power to harness it in their work. For this reason, it is important for staff and MPs to understand the basic tenets of data literacy and AI literacy, including the risks and downsides, before they utilize such tools in parliament.

Data literacy in an AI context

By establishing a training programme and building a solid foundation in data management, parliaments can harness the full potential of their data assets to support evidence-based decision-making, foster public trust and uphold democratic principles.

For further guidance on the training requirements for data literacy when using AI, refer to the sub-guideline Training for data literacy and AI literacy: Data literacy in an AI context.

Developing AI literacy

AI literacy is crucial in parliaments because it enables MPs, decision makers and staff to make informed choices about AI adoption, shape appropriate policies and regulations, and effectively oversee AI-driven initiatives.

Having a well-trained workforce that is familiar with at least the basic tenets of AI will help parliament to both leverage the opportunities AI presents for enhancing parliamentary functions and mitigate the potential risks that can occur.

For further guidance on developing AI literacy in parliaments, refer to the sub-guideline Training for data literacy and AI literacy: Developing AI literacy.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Training for data literacy and AI literacy: Data literacy in an AI context

About this sub-guideline

This sub-guideline is part of the guideline Training for data literacy and AI literacy. Refer to the main guideline for context and an overview.

Why data literacy matters

Data literacy is the ability to read, understand, create and communicate data as information. It involves understanding how to effectively collect, analyse, interpret and present data in meaningful ways. Data literacy includes knowing where data comes from, and grasping basic statistical concepts and data presentation and visualization techniques. Data literacy is vital for critically evaluating data-driven arguments and conclusions.

In parliaments, data has emerged as a critical asset across all business domains. As parliaments increasingly rely on data-driven insights to fulfil their mandate, the importance of robust data management practices cannot be overstated. Becoming data-literate empowers parliaments to harness the full potential of their data assets, driving informed decision-making, enhancing transparency, fostering public trust and strengthening the foundations of democratic governance. This is fundamental to the adoption of AI.

Data literacy training for MPs and non-technical staff

Users are important actors in a data culture, since they add data into many systems and are often the people who come face-to-face with the data. It is important for them to understand basic data principles and to be attuned to possible errors or problems that can arise.

Increasingly, too, users are extracting, combining and otherwise repurposing data, often into top-line reports or business dashboards. In all cases, this requires quality assurance, as well as an understanding of where the data comes from and what it means.

AI can be a powerful tool for analysing and understanding data and trends. But it can also be unreliable, which is why data literacy is especially important in this context.

Parliaments could develop or source training programmes on the following topics in order to achieve a good level of data literacy across MPs and a wide range of non-technical staff:

- Introduction to data literacy
- Data collection and management
- Data analysis fundamentals
- Data visualization and presentation
- Ethical data usage and privacy
- Data-driven decision-making
- Introduction to AI and machine learning
- Critical thinking with data

These courses could be offered at different levels (basic, intermediate or advanced) to cater to learners in different roles and with varying levels of exposure to AI-based systems within parliament.

Data literacy training for decision makers

AI-related data literacy is crucial for senior leaders and decision makers in parliaments, as it enables informed decision-making, effective risk management and strategic planning for AI adoption. Data-literate leaders are able to exercise oversight over AI projects, optimize resource allocation and foster innovation, while ensuring that AI is used ethically.

By equipping decision makers with these skills, parliaments can ensure that AI adoption is guided by informed leadership, aligning with institutional goals while adhering to ethical standards and best practices.

A targeted data literacy training programme for senior leaders and decision makers could include the following:

- **Dedicated workshop**: This session covers the importance of data-driven initiatives, AI readiness and associated risks. Participants gain an overview of foundational data management concepts tailored to their level, reaching a comprehensive understanding of the AI landscape in a parliamentary context.
- **Just-in-time learning and self-paced resources**: Additional learning resources allow senior leaders and decision makers to embed their knowledge and understanding of AI.

Data literacy training for technical staff

Data literacy training is of paramount importance for technical staff, as these individuals are at the forefront of implementing and managing AI systems in parliaments. Their expertise directly impacts the effectiveness, ethical use and security of AI applications in parliamentary operations.

A data literacy training programme for this population could be structured as follows, depending on the needs of parliament:

- **Data management foundations**: The programme begins with an overview of the fundamental concepts of data management, ensuring participants have a solid understanding of data types, integrity, governance and life-cycle management. This foundational knowledge is critical, since the quality and management of data directly affect the performance and reliability of AI systems.
- **Practical skills**: The programme then progresses through practical skills in data collection, cleansing and storage. These skills are essential for preparing and maintaining the high-quality data sets that AI systems rely on. Advanced topics such as database management, cloud storage solutions and data visualization techniques could be included to ensure technical staff can effectively handle and communicate insights from large, complex data sets.
- **Non-technical aspects**: Importantly, the programme also addresses key non-technical topics including data ethics, legal compliance and the application of data in parliamentary contexts. This training ensures that technical staff not only have the skills to implement AI systems, but also understand the broader implications and responsibilities of using AI in a parliamentary setting.

By providing comprehensive training in this way, parliaments can ensure that their technical staff are well-equipped to lead the responsible and effective implementation of AI technologies, ultimately enhancing the efficiency and effectiveness of parliamentary operations.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Training for data literacy and AI literacy: Developing AI literacy

About this sub-guideline

This sub-guideline is part of the guideline Training for data literacy and AI literacy. Refer to the main guideline for context and an overview.

Why AI literacy matters

AI literacy is the ability to understand, critically evaluate and effectively interact with AI technologies. In a parliamentary context, it encompasses knowledge of the capabilities and limitations of AI, as well as an understanding of its potential impacts on legislative processes and democratic governance. Moreover, AI literacy can help parliaments to foster a culture of responsible innovation.

AI literacy is crucial in parliaments because it enables MPs, senior parliamentary managers and staff to make informed choices about AI adoption, shape appropriate policies and regulations, and effectively oversee AI-driven initiatives.

Having a well-trained workforce that is familiar with at least the basic tenets of AI will help parliament to both leverage the opportunities AI presents for enhancing parliamentary functions and mitigate the potential risks that can occur.

AI literacy training is needed across a range of parliamentary stakeholders. The nature and content of this training will differ depending on their role. A potential outline of training programmes for MPs, decision makers and technical staff is given below.

AI literacy training for MPs

MPs face a dual challenge: they must understand AI in order to effectively use it and oversee its implementation within parliament, while also grasping its broader implications to inform their legislative work.

A flexible AI literacy programme, tailored to the unique culture of each parliament and the demanding schedules of its MPs, would be a useful way to address this need. A potential structure for such a programme is outlined below:

- **AI fundamentals**: The programme begins with a clear introduction to the fundamentals of AI, exploring its core concepts and potential applications in legislative work. MPs learn how AI can enhance parliamentary functions, from improving constituent services to streamlining research processes.
- **Risks and ethical implications of AI**: Since understanding the potential of AI is only half of the equation, the programme also delves into the associated risks and ethical implications, covering topics such as privacy concerns, potential biases and societal impacts.
- **AI governance**: This section of the programme covers AI governance and its role in risk mitigation. For parliaments with existing AI policies, the programme explains the relevance of these policies to MPs' daily work. For parliaments that already use AI systems in their legislative processes, the programme offers practical insights into the operation and impact of these systems.
- **Ethical principles**: Throughout the programme, emphasis is placed on the ethical principles that form the cornerstone of responsible AI use in democratic institutions. MPs explore concepts such as transparency, accountability and the preservation of human autonomy, all of which are crucial for promoting societal well-being in an AI-driven world.

By the end of this programme, MPs are equipped with the knowledge to confidently engage in AI-related policy discussions, make informed decisions about AI adoption in parliamentary processes, and navigate the increasingly AI-influenced landscape of modern governance. This comprehensive yet accessible approach ensures that MPs can harness the benefits of AI while safeguarding democratic values and protecting the public interest.

AI literacy training for decision makers

Parliamentary decision makers need a solid understanding of AI in order to guide their institutions effectively. They must equip themselves with essential knowledge about AI and its implications for parliamentary work. A possible structure for a training programme for this audience is given below:

- **AI fundamentals**: The training begins with an explanation of the fundamental concepts of AI and its potential applications within parliamentary contexts.
- **Governance, opportunities and risks**: Decision makers learn how AI can enhance legislative processes and streamline administrative tasks, as well as examining the potential risks of the technology. The programme also looks at how effective AI governance can mitigate these risks and ensure responsible use.
- **Broader implications**: The programme covers the broader legal, ethical and social implications of AI, covering key concepts such as privacy, transparency, accountability and fairness.
- **Limitations of AI**: Participants explore the limitations of AI – and of generative AI in particular – to ensure they are equipped with the knowledge needed for realistic implementation.

- **AI legislative frameworks and policies**: For parliaments with specific AI legislative frameworks or policies, the programme covers the implications of these frameworks or policies for parliamentary AI use cases.

AI literacy training for technical staff

Technical staff involved in AI production and procurement require a comprehensive understanding of AI technologies and their parliamentary context and application. A training programme for these staff could cover several key areas:

- **AI fundamentals:** The programme begins with an exploration of the history of AI, key concepts, and the role of the technology in governance and legislation. This foundational knowledge helps align AI implementations with parliamentary needs.
- **Machine learning and deep learning:** Participants explore various learning paradigms, common algorithms and frameworks. Hands-on training with relevant tools provides practical knowledge for developing AI models that address legislative challenges.
- **Ethical implications:** A significant portion of the programme focuses on the ethical and social aspects of AI. Participants engage in critical discussions about bias, fairness, transparency, and the impact of AI on privacy and human rights. This ethical grounding ensures AI systems uphold democratic principles.
- **Governance-specific applications:** Through case studies and practical projects, technical staff learn to identify opportunities for enhancing legislative processes with AI. They work on developing AI solutions for real parliamentary challenges, considering both technical feasibility and ethical implications.
- **AI tools and resources:** The programme introduces a range of AI development tools, platforms and resources. Participants gain experience with AI libraries, application programming interfaces (APIs), data sets and pre-trained models. They also practice creating effective prompts for generative AI tools while adhering to ethical principles.

For parliaments primarily using generative AI, the programme could be adapted to emphasize avoiding inaccuracies, hallucinations and biases in AI outputs. It could also stress the importance of clear guidelines to protect against adversarial prompts and to maintain information security.

By the end of this comprehensive programme, technical staff are well-equipped to lead AI initiatives within their parliaments. They possesses the technical expertise needed to implement AI solutions, the ethical foundation necessary for ensuring responsible use, and the contextual understanding required to align these technologies with parliamentary needs and values.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Planning and implementation

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Project portfolio management

Audience

This guideline is intended for parliamentary staff including senior leaders who may not have a technical background in AI but who are involved in the management or oversight of AI projects.

About this guideline

This guideline examines the practice of managing a portfolio of AI projects.

The ability to effectively manage a diverse array of AI projects is increasingly important as AI becomes more commonplace in parliaments. By strategically prioritizing and aligning AI projects with organizational goals, parliaments can maximize their impact.

This guideline therefore aims to equip parliaments with the knowledge and tools they need to navigate the complexities of AI project portfolio management, ensuring that AI technologies are implemented successfully and in line with ethical principles.

What is project portfolio management?

Project portfolio management (PPM) refers to the centralized management of a parliament's programmes, projects and related activities. It is used to meet strategic objectives by optimizing resource allocation, balancing risks and maximizing overall value.

Why project portfolio management is relevant to AI governance

AI project portfolio management (AI PPM) entails overseeing a diverse array of AI technologies and initiatives in order to optimize performance and achieve strategic objectives as part of ongoing strategic governance.

One key consideration in AI PPM is recognizing that AI systems are designed to evolve over time, adapting to changing needs and circumstances. With multiple large

language models (LLMs) available in the field of generative AI, parliaments must carefully evaluate and select the most suitable AI technologies to support their objectives.

Effective AI PPM involves ensuring that AI initiatives are aligned with organizational goals, while also prioritizing transparency, accountability and ethical considerations. By adopting a proactive approach to AI PPM, parliaments can harness the full potential of AI to drive innovation, efficiency and effectiveness in legislative processes and governance.

From segmentation to portfolio management

AI PPM provides a structured approach to strategically selecting, prioritizing and overseeing AI projects within parliaments. Unlike traditional project management, which focuses on individual projects, AI PPM takes a holistic view, considering the collective impact of AI projects and their alignment with organizational goals and priorities.

At its core, AI PPM involves identifying, evaluating and prioritizing AI projects based on their potential value, feasibility and alignment with strategic objectives. Each potential or existing AI project can be evaluated against seven key criteria:

1. **Strategic alignment**: Consider whether the project aligns with parliament's overarching strategy and goals. Projects that contribute directly to achieving strategic objectives should be given higher priority.
2. **Measurable impact**: Prioritize projects with clear, objective and measurable impact metrics in order to ensure a tangible return on investment.
3. **Augment or replace**: Determine whether the project will augment current human operations or replace existing manual processes entirely. Projects that enhance human capabilities or efficiency should be prioritized over those that solely aim to automate existing processes.
4. **Nature of the problem**: Evaluate whether the problem being addressed is suitable for AI-driven solutions. Projects that align with the capabilities of AI technologies and have clear problem-solving potential should be prioritized.
5. **Data availability**: Assess the availability and quality of data required for the project. Projects for which the necessary data is already available or can easily be acquired should be prioritized, as data availability is essential for AI model training and performance.
6. **Technological capability and skills**: Consider whether parliament possesses the technological infrastructure and skill set required to develop, deploy and scale up the solution successfully. Projects that align with existing technological capabilities and expertise should be prioritized in order to minimize implementation challenges.
7. **Ethical considerations**: Finally, ensure that all ethical considerations – including bias mitigation, privacy protection and transparency – have been thoroughly evaluated for each project and that the project aligns with agreed organizational values.

Prioritization

Once parliament has identified and evaluated AI projects, the next important step is to review the institution's (potential) AI portfolio as a whole and to prioritize the workstream:

- Evaluate how each AI project aligns with parliament's overall strategic goals and rank them according to potential value and impact.
- Assess resource availability, conflicts and constraints.
- Evaluate the potential benefits and risks of each AI project and prioritize those with favourable risk-reward ratios.
- Identify projects that are prerequisites for others or that could create synergies if implemented together, and consider prioritizing those that unlock value in other projects or create a foundation for future initiatives.
- Understand the time sensitivity of each project, looking to balance quick wins and long-term strategic value.
- Assess the impact of each project on key stakeholders (both internal and external) and prioritize those with high levels of stakeholder support.

Methodologies and frameworks for PPM

There are many standards and frameworks that organizations can use to support the implementation of a PPM approach. Some examples are given below:

- The Standard for Portfolio Management (SPM): a standard developed by the Project Management Institute (PMI)
- Disciplined Agile (DA): a toolkit, also developed by PMI, that includes portfolio management practices
- https://www.pmi.org/learning/library/pathway-organizational-project-management-maturity-8221Organizational Project Management Maturity Model (OPM3): another model developed by PMI
- Projects IN Controlled Environments (PRINCE2): a project management method that also has implications for portfolio management
- Lean Portfolio Management: a method that is part of the Lean-Agile approach
- Hoshin Kanri: a strategic planning process that can be applied to portfolio management
- Objectives and Key Results (OKRs): an approach that can be used to align portfolios with organizational goals

Other approaches that can be used for PPM include Balanced Scorecard and Theory of Constraints.

Parliaments can use their existing methodologies to support AI PPM, or they can adopt an established external framework that fits well with their culture and working methods.

For further guidance on developing or adopting a framework for PPM, refer to the sub-guideline Project portfolio management: The STEP approach.

Find out more

- IPU Centre for Innovation in Parliament (CIP), IT Governance Hub: Framing the development of IT governance for parliaments

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Project portfolio management: The STEP approach

About this sub-guideline

This sub-guideline is part of the guideline Project portfolio management. The main guideline should be read first for context and an overview.

This sub-guideline introduces an example of a project portfolio management (PPM) methodology that parliaments could implement to support their AI programme.

The STEP approach

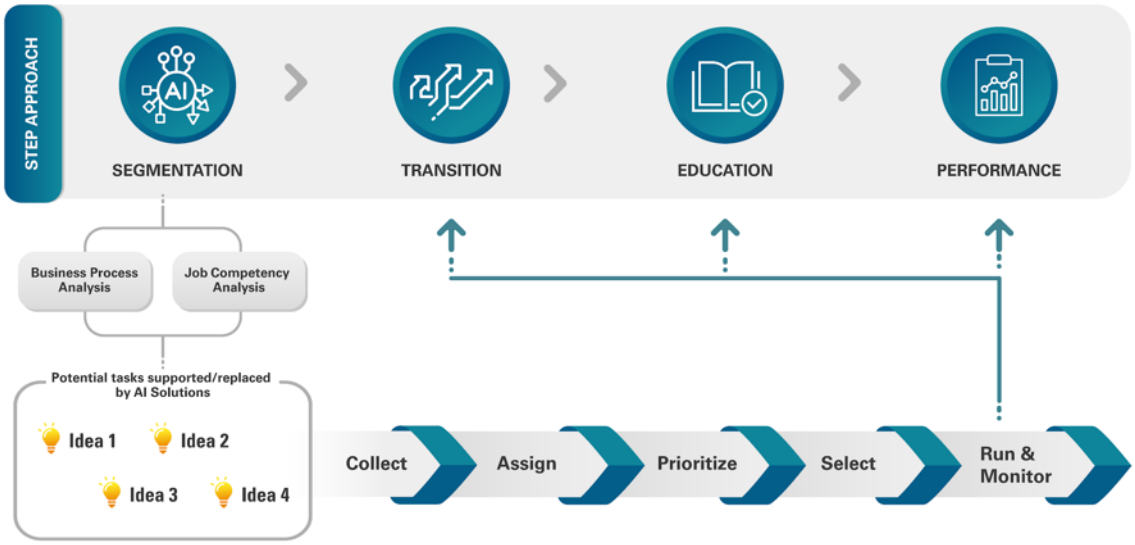
The STEP approach offers a structured way for parliaments to implement AI technologies and leverage their potential to enhance productivity and efficiency. Following this model allows parliaments to navigate the introduction of AI solutions with precision and foresight, ultimately driving positive outcomes and organizational success.

The STEP approach comprises four key stages:

- **Segmentation** involves identifying and segmenting tasks suitable for AI integration, recognizing that no single AI system can fulfil all requirements.
- **Transition** is the process of gradually incorporating AI systems into existing workflows, allowing for the deepening or upgrading of work roles in order to maximize the benefits of AI technologies.
- **Education** plays a crucial role in ensuring staff readiness and competency in effective AI use, encompassing training on AI functionality, potential biases and best practices for integration.
- **Performance evaluation** measures the impact and efficacy of AI systems in real-world contexts, providing valuable insights into their tangible benefits and areas for improvement.

These stages are shown Figure 1 and discussed in turn below.

Figure 1: The STEP approach



Segmentation

Segmentation offers a systematic framework for analysing activities within parliamentary proceedings by delineating related processes and tasks.

Segmentation entails categorizing tasks into three distinct types:

- Tasks that AI cannot or should not perform according to decisions made by those in charge of AI governance
- Tasks where AI can augment staff actions (augmentation)
- Tasks that can be automated by AI (automation)

Crucially, this approach emphasizes that parliamentary staff should take the lead in task segmentation, leveraging their domain expertise to identify suitable tasks for AI integration. It also underscores the importance of staff experimentation with AI tools before widespread adoption, enabling them to assess suitability, usability and effectiveness in real-world contexts.

By adopting a proactive approach to task segmentation following a business analysis or a job competency analysis, parliaments can maximize the benefits of AI systems while ensuring alignment with organizational goals and priorities. Having a clear list of tasks identified for augmentation or automation will be crucial for the ideation phase. Within an AI PPM context, these tasks become the primary candidates for digital transformation.

Transition

The transition phase of the STEP approach focuses on changes to parliamentary staff members' roles as tasks are augmented or automated through the integration of AI systems.

During this phase, resources freed up by AI implementation are carefully considered and assigned to one of three categories:

- Job elimination: resources are redeployed to other functions within parliament
- Job upgrading: tasks previously performed by more senior staff are now supported by AI-enabled systems
- Job deepening: staff members have the opportunity to spend their time on tasks with greater added value, gain deeper knowledge in their domain, or engage in skill-building activities such as preparing training materials

By strategically managing this transition, parliaments can optimize resource allocation, enhance staff productivity and capitalize on the transformative potential of AI technologies.

Education

In the context of parliamentary AI integration, education is a multifaceted approach aimed at fostering a culture of continuous learning and skill development, with a particular focus on embedding an ethos of AI learning into the organizational culture. As part of this, parliaments should frequently review and update learning materials to ensure they keep pace with evolving AI trends and best practices, as well as revisiting the segmentation process regularly to identify new opportunities for AI integration.

Parliamentary staff can be enrolled on certifying courses that teach them the necessary knowledge and skills to effectively leverage AI solutions in their roles. These skills could include fine-tuning documents or data priority from the organization, mastering prompt engineering to create effective commands or prompts for AI systems, and evaluating the validity of predictions made by these systems.

Performance evaluation

The performance evaluation phase acknowledges that the introduction of AI systems in parliaments represents a shift rather than just a lift in operational dynamics. Performance evaluation encompasses the following aspects:

- Recalibrating individual annual performance metrics to encompass factors such as speed, efficiency, accuracy and creativity, reflecting the contribution of AI systems and related training
- Identifying and measuring appropriate performance indicators for AI systems to determine whether these systems are having a measurable impact that aligns with the organization’s strategic objectives
- Regularly assessing the viability and effectiveness of AI systems, with feedback loops that provide valuable qualitative metrics to evaluate adoption rates and user satisfaction
- Evaluating the effectiveness of AI-related training programmes, and continuously updating learning materials to ensure staff remain up to date with the latest AI advancements and methodologies
- Closely monitoring the availability of new AI technologies and reviewing task segmentation accordingly

Summary

The STEP approach provides a structured and adaptable AI PPM method for parliaments. It helps ensure alignment with organizational goals, optimizing resource allocation and maximizing the overall impact of AI investments. By adopting this approach, parliaments can navigate the complex landscape of AI projects with confidence, driving innovation, efficiency and effectiveness in legislative processes and governance.

Find out more

- IPU Centre for Innovation in Parliament (CIP), IT Governance Hub: Framing the development of IT governance for parliaments

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Data governance

Audience

This guideline is intended for parliamentary staff involved in the oversight and implementation of AI-based systems, including business managers, chief information officers, chief technology officers and IT managers. It will also be useful for senior parliamentary managers responsible for AI governance.

About this guideline

This guideline outlines the desirable characteristics of data for safe AI systems development. It highlights data-quality issues to avoid during the planning and development of AI systems, and explains how to implement privacy principles to protect personal data – an essential factor for AI development. These practices should be guided by data governance that aligns with the organization's needs.

Since data exists independently of AI systems and is used by parliaments, data management practices should be established before any AI initiatives commence. This lays the foundation for trustworthy AI systems in legislative bodies.

Why is data governance relevant to AI governance?

The quality and reliability of AI system outputs depend heavily on two factors:

- The quality of data used to train the AI model
- The quality of data the AI system uses during its operation

It is crucial to protect this data from unauthorized access and misuse. Improving data quality and enhancing data protection are key components of an organization's AI governance strategy. Achieving these improvements requires coordinated actions and agreements among various stakeholders involved in data-related decisions and processes.

Data governance plays a crucial role here. It involves coordinating and managing the efforts of all stakeholders to enhance data quality and protect privacy. By effectively implementing data governance practices, organizations can ensure they are developing AI systems that are reliable and trustworthy.

Data quality

Data quality refers to certain features of data that make it accessible, useful and reliable to support effective decision-making. Data must be demonstrably accurate, complete, consistent, accessible, relevant and secure.

For further discussion of this subject, refer to the sub-guideline Data governance: Data quality.

Personal data protection

Personal data protection refers to practices, policies and legislation designed to safeguard individuals' personal data from unauthorized access, misuse or exposure. It encompasses various measures to ensure that personal data is collected, stored, processed and shared in a way that respects individuals' privacy and complies with relevant laws and regulations.

For further discussion of this subject, refer to the sub-guideline Data governance: Personal data protection.

Data governance in a parliamentary context

Data governance is crucial for coordinating efforts to enhance data quality and privacy. It encompasses policies, roles, responsibilities, processes and technology aimed at improving data quality and creating an optimal data environment.

For further discussion of this subject, refer to the sub-guideline Data governance: Data governance in a parliamentary context.

Data management for AI systems

Data management is a crucial factor in the implementation of AI systems in parliaments. Staff must understand the key steps for establishing effective data management practices, including creating governance policies, managing metadata, ensuring data quality and protecting personal information.

For further discussion of this subject, refer to the sub-guideline Data governance: Data management for AI systems.

Find out more

- Australian Government, Office of the Australian Information Commissioner: Australian Privacy Principles guidelines
- Government of Brazil: Guia de Elaboração de Programa de Governança em Privacidade (available in Portuguese only)
- Government of the United Kingdom: The Government Data Quality Framework
- Government of the United Kingdom: Using personal data in your business or other organisation
- Norwegian Data Protection Authority: Artificial intelligence and privacy

- Publications Office of the European Union: Data quality requirements for inclusive, non-biased and trustworthy AI: Putting science into standards

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Data governance: Data quality

About this sub-guideline

This sub-guideline is part of the guideline Data governance. Refer to the main guideline for context and an overview.

This sub-guideline explains why data quality matters, explores the dimensions of data quality, and examines both the benefits of high-quality data and the risks associated with low-quality data.

Why data quality matters

Data quality refers to certain features of data that make it accessible, useful and reliable to support effective decision-making. Data must be demonstrably accurate, complete, consistent, accessible, relevant and secure.

High-quality data is essential for parliamentary processes. It forms the foundation of reliable, data-driven decisions that can improve operational efficiency, reduce both operational and strategic risks, and help meet overall business needs. When AI systems are integrated into decision-making processes, they can amplify the benefits of digital solutions. However, it is important to note that this integration can also magnify any associated risks. Therefore, ensuring data quality becomes even more critical when AI is involved in parliamentary operations.

Dimensions of data quality

Data quality can be evaluated according to a number of criteria, such as the following:

- **Accessibility** is the extent to which information is available, or easily and quickly retrievable. Data availability requirements per business process should be identified.
- **Appropriate amount of information** is the extent to which the volume of information is appropriate for the task at hand. This can be managed by defining, for each data element, how critical the amount of information captured is for analysis purposes.

- **Completeness** is the extent to which information is not missing and is of sufficient breadth and depth for the task at hand. This aspect is managed by reporting on the completeness of data fields and devising plans to capture all data as per business requirements.
- **Consistency** is the extent to which information is presented with the same content across multiple systems and platforms. This aspect is managed by understanding data standards and business rules and ensuring that systems adhere to the defined rules.
- **Freedom from errors** is the extent to which information is correct and reliable, and the degree of agreement between a data value (or set of values) and a source assumed to be correct. Freedom from errors may be attained by defining validation rules, conducting regular testing and reporting on samples of data for compliance.
- **Relevance** is the extent to which information is pertinent to business processes. This aspect is managed by determining the business use of each data element and assessing its value and relevance through user feedback.

Benefits of high-quality data

High-quality data has numerous benefits for parliaments, including the following:

- **Trustworthy AI systems**: Accurate and reliable data enables better-trained AI models.
- **Improved analytics**: High-quality data enhances the accuracy of analytics, leading to more precise and actionable insights, conclusions and predictions.
- **Improved decision-making**: Accurate and reliable data allows parliaments to make better-informed decisions with less risk.
- **Reduced risk**: Secure and protected data can help to prevent fraud, financial losses and reputational damage.
- **Cost savings**: High-quality data reduces the costs associated with correcting errors, and reduces the time needed to deal with data-related issues.
- **Informed policymaking**: Accurate and reliable data enables lawmakers to draft bills based on solid evidence, leading to more effective and impactful legislation.
- **Support for legislative research**: High-quality data is crucial for conducting thorough legislative research, enabling MPs to understand complex issues and make informed decisions.
- **Support for innovation**: High-quality data can be used to develop new products and services, since the data can reveal not only hidden problems but also potential ways to solve such problems.
- **Enhanced public trust**: Transparent and high-quality data fosters trust among citizens, who can see that decisions are based on accurate information.

Key data-quality issues

The list below outlines some of the primary obstacles to high-quality data:

- **Data integrity issues**: Errors in data entry or processing can compromise the accuracy and reliability of data.

- **Duplicate data**: Having multiple records for the same entity can lead to confusion and errors in reporting and analysis, making it difficult to keep different versions of data in sync across operations.
- **Inconsistent data**: Variations in data formats or standards across different systems can cause integration issues and inaccuracies.
- **Outdated data**: Using old or obsolete data can lead to decision-making based on irrelevant information.
- **Incomplete data**: Missing data fields can lead to gaps in analysis and hinder comprehensive decision-making.
- **Ambiguous data**: Data that is unclear or lacks context can be misinterpreted, leading to incorrect conclusions.
- **Misinformed decisions**: Inaccurate or incomplete data can lead decision makers to make the wrong choices.

Risks associated with low-quality data

In parliamentary contexts, low-quality data can cause a unique set of problems, some of which are discussed below:

- **Misguided legislation**: When flawed data is used for evidence, research or policy analysis, this can lead to legislation that is ineffective or has unintended negative consequences. For example, an environmental protection bill based on inaccurate pollution statistics might target the wrong industries or fail to address the most pressing issues.
- **Ineffective resource allocation**: Budgetary decisions based on unreliable data can lead to inefficient resource allocation. For example, allocating funds to a social programme based on outdated poverty statistics might result in the programme not reaching the communities most in need.
- **Reduced public trust**: The perception that parliamentary decisions are based on questionable data can erode trust in the legislative process and undermine the effectiveness of government institutions.
- **Data privacy concerns**: Mishandling sensitive data can lead to breaches and loss of public trust. For example, a leak of citizens' personal data can cause a public outcry and lead to legal repercussions.
- **Biases in data collection**: Inherent biases in data-collection methods can skew policy outcomes. For example, surveys that do not adequately represent minority groups can lead to policies that overlook their needs.
- **Overreliance on quantitative data**: Ignoring qualitative data leads to incomplete analysis. For example, focusing solely on statistical data without considering public opinion or sentiment can result in unpopular and ineffective policies and mean that key social trends or attitudes are overlooked.

For example: if a parliamentary committee uses flawed crime statistics to justify increased police funding, the subsequent revelation of data errors could damage public confidence in both the decision-making process and the resulting policy changes.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Data governance: Personal data protection

About this sub-guideline

This sub-guideline is part of the guideline Data governance. Refer to the main guideline for context and an overview.

This sub-guideline looks at data protection regulations, specific types of sensitive data, and broader data privacy issues in the context of AI systems.

What is personal data protection?

Personal data protection refers to practices, policies and legislation designed to safeguard individuals' personal data from unauthorized access, misuse or exposure. It encompasses various measures to ensure that personal data is collected, stored, processed and shared in a way that respects individuals' privacy and complies with relevant laws and regulations.

Data protection regulations

Most countries have data protection regulations that govern how personal data is used. In all cases, parliaments must comply with relevant local regulation(s) by adopting or adapting the measures and processes required by law.

While the exact rules will vary in each case, such regulations generally impose the following requirements:

- Personal data must be used fairly, lawfully and transparently.
- Personal data must be used for specified, explicit purposes.
- Personal data must be used in a way that is adequate, relevant and limited to only what is necessary.
- Personal data must be accurate and, where necessary, kept up to date.
- Personal data must be retained for no longer than is necessary.
- Personal data must be handled in a way that ensures appropriate security, including protection against unlawful or unauthorized processing, access, loss, destruction or damage.

Sensitive data

There may be stronger legal protections for more sensitive information, such as the following:

- Race and/or ethnic background
- Political opinions
- Religious beliefs
- Trade union membership
- Genetics
- Biometrics (where used for identification)
- Health
- Gender
- Sexual orientation

Data privacy issues

In today's increasingly digitized society, there is a growing risk of data being wrongly shared, stolen or leaked, and of inaccuracies perpetuating through multiple systems. In the context of AI, some of the issues that must be addressed include the following:

- **Exposure to privacy breaches and security incidents**: Data breaches might cause parliament to suffer long-lasting reputational damage and legal consequences, including fines, lawsuits and other regulatory sanctions.
- **Overcollection and mismanagement of data**: Collecting more data than necessary can heighten the risk of breaches and privacy violations, as well as increase the complexity of data-management processes.
- **Bias**: The use of AI can introduce biases into decision-making processes, leading to unfair treatment of, and discrimination against, individuals based on their data.
- **Intrusive surveillance**: When data is used unethically for intrusive surveillance of individuals' personal life or for behaviour profiling, parliament runs the risk of both legal action and reputational damage.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Data governance: Parliamentary context

About this sub-guideline

This sub-guideline is part of the guideline Data governance. Refer to the main guideline for context and an overview.

This sub-guideline examines the importance of data governance and provides an overview of related roles and responsibilities within parliament.

Why data governance matters

Data governance is crucial for coordinating efforts to enhance data quality and privacy. It encompasses policies, roles, responsibilities, processes and technology aimed at improving data quality and creating an optimal data environment.

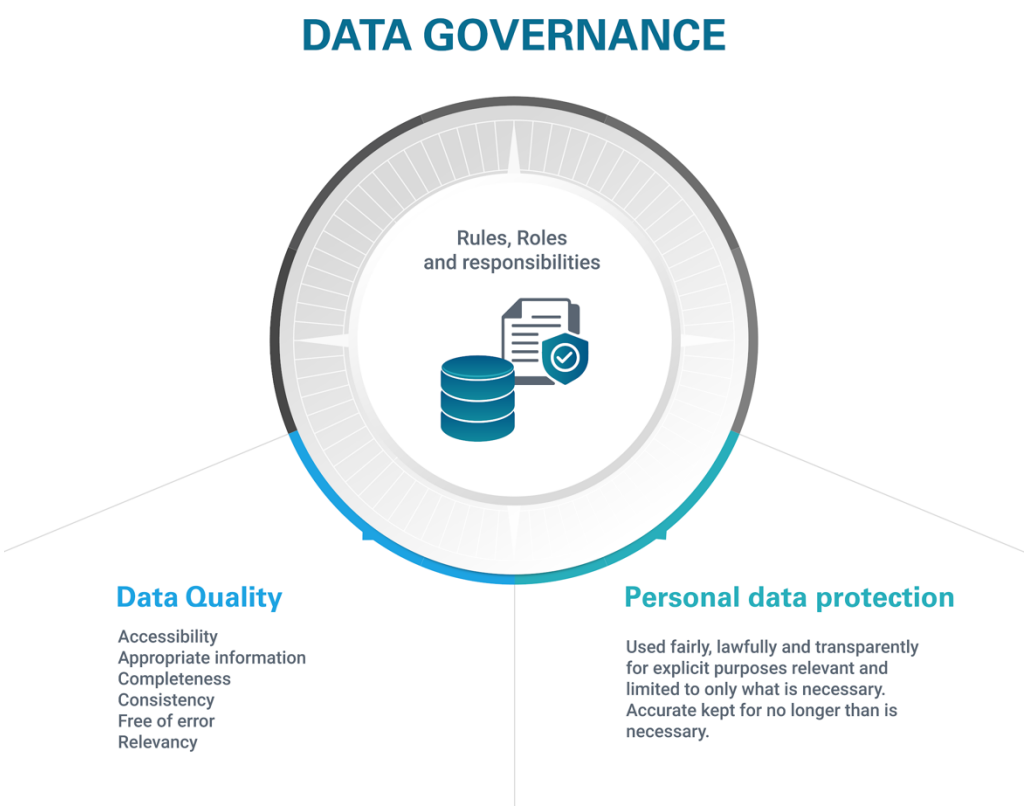
Data governance provides guidance, assurance and support to transform parliaments into data-driven organizations through the following measures:

- Clearly defining roles and responsibilities for data management to ensure efficiency
- Introducing common rules, guidelines and practices for consistent data management across the organization
- Ensuring data security and protecting privacy
- Promoting easy access to, and usability of, data to support informed decision-making
- Maintaining the accuracy and integrity of data throughout its life cycle
- Ensuring compliance with relevant laws and regulations in data management practices
- Providing high-quality data to inform policymaking and strategic planning
- Fostering a data-centric culture that emphasizes the importance of privacy, data quality and data governance across all parliamentary units and processes

Roles and responsibilities in data governance

Figure 1 provides an overview of roles and responsibilities in data governance. Further discussion can be found in the remainder of this section.

Figure 1: Roles and responsibilities in data governance



Data owner

Every piece of data should have a data owner, who is in charge of decisions regarding data protection, data storage, data classification, data access, data formats, metadata, and all improvements necessary to make the data useful for parliament's needs.

Data steward

Data owners designate a data steward to support them with managing data quality, ensuring compliance with policies, and facilitating communication between data owners and users.

Establishing data ownership

Practices for establishing data ownership can include the following:

- Establish clear ownership roles, assigning specific individuals or teams as data owners for each data set. These owners are responsible for the accuracy, integrity and security of the data throughout its life cycle.
- Establish data stewardship roles to support data owners.

- Create ownership policies that outline the responsibilities and expectations of data owners, including practices for data creation, maintenance, access and disposal.
- Implement role-based access controls to ensure that only authorized individuals can access and modify data.
- Conduct regular audits to verify that data ownership policies are being followed and that data quality standards are maintained. Use audit findings to make continuous improvements.
- Provide training programmes for data owners and stewards to ensure they understand their roles and responsibilities.
- Encourage collaboration between data owners, data stewards and users to ensure that data management practices are aligned with parliament's goals. Establish forums or committees for discussing data-related issues and making collective decisions.
- Implement monitoring tools to track data quality and ownership compliance. Generate regular reports to keep stakeholders informed of the status of data governance initiatives.
- Define data life cycle management practices, covering the entire process from creation to disposal. Ensure that data owners are responsible for overseeing this process.

Responsibilities of data owners

The responsibilities of data owners are as follows:

- Appoint one or more data stewards to support them in their role and to facilitate the implementation of the data ownership guidelines.
- Allocate resources to information ownership objectives.
- Apply operational guidelines and procedures for information ownership.
- Establish and measure data performance metrics and communicate about actual data performance.
- Establish future data requirements based on strategies and business trends.
- Position and manage data as a corporate asset.

Responsibilities of data stewards

The responsibilities of data stewards – some from business units, others from the IT unit – are as follows:

- Deploy and implement information ownership programmes, operational guidelines and procedures.
- Support audits of information ownership processes.
- Upgrade and develop information ownership processes.
- Provide intensive data-quality training to parliamentary staff.
- Help parliamentary staff to identify and solve data-related issues and problems.
- Determine and interpret trends in data quality.
- Support data-quality improvement efforts.
- Serve as first-line support to help users solve data accuracy issues, data definition issues and data usage issues.

Responsibilities of the chief information officer (or equivalent)

Parliament's chief information officer (or equivalent), acting as the data custodian, has overall responsibility for the following aspects:

- Implement and maintain the infrastructure needed to deliver data from its point of capture and storage to a point of need.
- Manage the availability of systems to access, retrieve and manipulate data.
- Ensure the integration and consistency of data across multiple applications and sources.
- Ensure that backup and recovery procedures are in place to prevent data loss.
- Secure data access and provide up-to-date solutions to protect against malicious code.
- Implement an IT helpdesk function for the following purposes:
    - Logging data issues
    - Escalating data issues to the appropriate information owner
    - Connecting users with second-line support to help interpret data fields and/or information within data fields
    - Granting controlled access to data to authorized users
    - Optimizing operational efficiency and effectiveness
    - Monitoring and reporting data issues

Responsibilities of parliamentary users

Parliamentary users of AI-based systems have the following responsibilities:

- Participate in data-improvement initiatives.
- Enhance their job skills related to improved data quality.
- Capture data and utilize it in parliament's processes in order to optimize operational efficiency and effectiveness.
- Monitor and report data issues to the data owner and/or the chief information officer.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Data governance: Data management for AI systems

About this sub-guideline

This sub-guideline is part of the guideline Data governance. Refer to the main guideline for context and an overview.

This sub-guideline focuses on effective data management, which is a crucial aspect of implementing AI systems in parliaments. It emphasizes that staff must understand the key steps for establishing effective data management practices, including creating governance policies, managing metadata, ensuring data quality and protecting personal information. By following these recommendations, parliaments can build a solid foundation for trustworthy AI systems, enhancing their ability to make data-driven decisions and operate efficiently.

Introduction

Parliaments that are planning to implement data management for AI systems should consider the following prerequisites, which are explained in more detail below:

- Establishing a corporate data governance programme
- Establishing a data governance policy
- Establishing a communication plan
- Implementing a metadata management process
- Implementing a data quality process
- Implementing a personal data protection process

Establishing a corporate data governance programme

In order to put in place an institutional data management programme to leverage AI systems, parliaments must establish a corporate data governance programme with clear roles and responsibilities, and adhere to general rules for executing data management processes. A corporate approach, aligned with parliament's AI strategy, is the most effective way to ensure that data within future AI systems is compliant and managed effectively.

Establishing a data governance policy

The first concern in any data governance implementation is to develop and publish a data governance policy that outline roles and responsibilities and determines the rules for the corporate data management processes.

In general, a data governance policy covers the following aspects:

- The organizational structure related to data governance
- The roles that will interact with data governance processes
- The competencies or job descriptions expected for each role
- The relevant data governance processes

Establishing a communication plan

Once the data governance policy in place, the next recommended step is to draw up a communication matrix showing, in detail, the interactions determined by the policy, encompassing the main roles and their responsibilities.

Implementing a metadata management process

A metadata management process allows parliament to become acquainted with its own data assets, which is crucial for data governance.

Parliaments should identify relevant information to be captured as metadata – based on the goals established by the organization's corporate strategy – as well as the specific information considered useful for data management and data description.

Below are some examples of metadata that could be captured for parliament's corporate data assets:

- Title (e.g. name of the bill or legislation + issue date)
- Description (e.g. date on which the bill or legislation was presented to the parliamentary board)
- Data owner (e.g. "Secretary of the Parliamentary Board")
- Data steward (e.g. "Protocol Registration Officer")
- Date format (e.g. "dd/mm/yyyy")
- Information systems (e.g. "Protocol Registration System")
- Main source (e.g.: "LegislationBills_DB")
- Personal data ("Yes"/"No")
- Sensitive data ("Yes"/"No")

As Figure 1 below shows, having a clearly identified metadata repository is crucial for understanding aspects such as the following:

- What the correct meaning of each data item is
- Who the data owner is
- Who the data steward is
- Whether the data is sensitive
- Which processes depend on the data

Figure 1: Structure of a metadata repository

Parliaments are advised to undertake continuous maintenance activities – such as frequent metadata review, validation and updating – in order to ensure that the metadata is precise, consistent and up to date.

Implementing a data quality process

The purpose of a data quality process is to ensure that data is managed in accordance with the rules laid down in the data governance policy. The main activities in this process are as follows:

- Data profiling:
  - Analysing data structure and contents
  - Identifying patterns, inconsistencies and anomalies in data
- Data quality requirements definition:
  - Establishing quality metrics and criteria (precision, completeness, consistency, uniqueness)
  - Defining business standards and rules to guarantee data compliance
- Data validation:
  - Applying business rules to validate data precision and consistency
  - Verifying whether the data meets the defined requirements
- Data cleansing/correction:
  - Fixing or removing incorrect, incomplete or duplicated data
  - Standardizing data formats
- Data integration:
  - Combining data from different sources and ensuring it remains consistent and correct
  - Solving data conflicts and eliminating duplications
- Data enrichment:
  - Incorporating additional information in order to increase data usefulness and completeness
- Data-quality monitoring:
  - Implementing continuous processes to monitor data quality
  - Using dashboards and reports to track data quality rates

Implementing a personal data protection process

The purpose of a personal data protection process is to ensure that parliament complies with privacy and data protection regulations, giving data subjects the necessary confidence to trust the institution with their personal data. The process dictates and influences how personal data is handled throughout its entire life cycle, encompassing the relevant strategies, skills, people, processes and tools.

The main steps in implementing a personal data protection process are as follows:

- Appointing a data protection officer
- Aligning the process with the expectations of senior parliamentary managers
- Assessing the maturity of parliament's existing corporate data protection arrangements
- Adopting data security measures to raise this level of maturity
- Establishing an organizational structure for the governance of personal data protection
- Implementing a personal data inventory
- Reviewing contracts related to the processing of personal data
- Preparing a personal data protection impact report
- Establishing terms and conditions for personal data protection
- Implementing an incident management process

Formalizing existing governance processes

Parliaments that already use data will, to some degree, have existing data governance and data management processes in place. Rather than creating a burdensome list of new responsibilities for business stakeholders, it is advisable to try to match AI-related tasks to existing job routines.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Security management

Audience

This guideline is intended for senior technical staff involved in the development and implementation of AI systems. Some of the material it contains may also be relevant to senior parliamentary managers looking to gain a better understanding of technical issues relating to security.

About this guideline

This guideline addresses the protection of AI systems from a wide range of threats and risks, outlining a set of practices for ensuring the confidentiality, integrity and availability of AI systems and data in parliaments.

AI systems are being deployed in many different ways, always with the aim of helping professionals increase their productivity. Parliaments are undoubtedly going to follow this trend, aiming for faster processes that accelerate democracy without harming or hurrying debate. Processes are also expected to be also safer, since parliaments are potentials target for national and international interest groups.

Why security management matters

As the use of AI increases, so does the risk to organizations using this technology. AI systems are associated with a range of security issues, such the inference of data – sometimes sensitive data – used in the training process, the alteration of such data, and the use of a particular prompt – wilfully or not – that could lead the AI system to reach a wrong or unexpected conclusion. All of these issues and more must be addressed before the AI system can be handed over to users.

Moreover, some AI behaviours could have a significant negative impact on an organization's public reputation. This means that AI systems can only be deployed after – at the very least – a basic risk assessment demonstrating that risks are low or controlled, and that the benefits outweigh these risks.

The deeper the knowledge someone has about AI, the easier it will be for this person to come up with a possible way of misleading the system and turning a breakthrough technology into a personal weapon to threaten different actors.

Moreover, even organizations that do not use AI models and systems are at risk, because criminals are already using AI in an attempt to increase the success rate of

their attacks. However, security considerations are especially important for organizations that do use AI in their own systems, since these models are prone to new types of attacks.

Considering the rise in cyberattacks, which surged after the COVID-19 pandemic, and the increasing use of AI models, which are the new “holy grail” of technology, overcoming AI threats is an important part of an organization’s cybersecurity plan.

Cybersecurity management in a parliamentary context

Cyberattacks are a growing concern as parliaments increase their reliance on internet-enabled connectivity – whether cloud-based servers, external systems or for users. Effective cybersecurity management is therefore critical for avoiding such attacks or minimizing their impact.
For further discussion of this subject, refer to the sub-guideline Security management: Parliamentary context.

Cybersecurity threats to AI systems

AI systems learn from the data they are fed and then apply models to help them make decisions, generate new content or do anything else they are programmed to do.
Just as parliaments must ensure that data is valid and of high quality, they must also ensure that there are no opportunities for attackers to exploit inputs into AI systems in order to corrupt and manipulate the data, modelling and outputs.

Attacks can occur in any phase, from data preparation through to AI system development, deployment and operation (for further discussion of this subject, refer to the guideline Systems development). As a result, the entire AI system life cycle should be properly supervised in order to minimize unexpected behaviours.
For further discussion of this subject, refer to the sub-guideline Security management: Threats.

Good practices for implementing AI-focused cybersecurity

Most types of attacks can be avoided or minimized by implementing good practices. Nonetheless, some attacks specifically targeting AI systems require specific countermeasures.

For further discussion of this subject, refer to the sub-guideline Security management: Good practices.

Main considerations when implementing AI-focused cybersecurity controls

Based on the main security frameworks, parliaments should gradually implement controls in the following four areas, according to their specific structure, needs and threat risks:

- Technical controls
- Organizational controls

- Human controls
- Physical controls

Together, measures across these four areas enable parliaments to enhance the protection of their AI systems.

For further discussion of this subject, refer to the sub-guideline Security management: Implementing cybersecurity controls.

Find out more

- Athalye A., N. Carlini and D. Wagner: Obfuscated Gradients Give a False Sense of Security: Circumventing Defenses to Adversarial Examples
- Deloitte: Impact of COVID-19 on Cybersecurity
- Federal Bureau of Investigation (FBI): FBI Warns of Increasing Threat of Cyber Criminals Utilizing Artificial Intelligence
- Hurst A.: NCSC releases guidelines for secure AI development
- International Monetary Fund (IMF): Rising Cyber Threats Pose Serious Concerns for Financial Stability
- Kurakin A., I. Goodfellow and S. Bengio: Adversarial Examples in the Physical World
- Kurakin A. and others: 9 Common Types of Attacks on AI Systems
- Norwegian National Security Authority (NSM): NSM ICT Security Principles
- Open Worldwide Application Security Project (OWASP): OWASP Machine Learning Security Top Ten
- Oseni A. and others: Security and Privacy for Artificial Intelligence: Opportunities and Challenges
- Saleous A. and others: COVID-19 pandemic and the cyberthreat landscape: Research challenges and opportunities
- Schneider S. and others: Designing Secure AI-based Systems: a Multi-Vocal Literature Review

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Security management: Parliamentary context

About this sub-guideline

This sub-guideline is part of the guideline Security management. Refer to the main guideline for context and an overview.

Why cybersecurity is important for AI

Parliaments' growing reliance on internet-connected systems, including AI, increases existing risks and introduces new ones, ranging from the manipulation of legislative processes to the disruption of administrative tasks. For instance, cyberattacks could alter the content of bills, influence voting outcomes or compromise AI-driven productivity tools. As parliaments adopt AI to manage public participation at scale, these systems may become targets for groups seeking to manipulate democratic processes.

Effective cybersecurity strategies and management practices help to maintain a secure digital environment that protects the integrity of parliamentary operations – especially as AI usage continues to grow. It is key to mitigating risks, ensuring that AI systems enhance rather than compromise parliamentary functions.

AI-related security considerations for parliaments

For AI systems, risks extend beyond traditional concerns, requiring specific consideration:

- **Integrity of the legislative process**: Cyberattacks could manipulate AI systems to alter bills, skew voting results or introduce biased information, potentially undermining the democratic process.
- **Administrative efficiency**: While AI can enhance productivity in administrative tasks, it also introduces new vulnerabilities. Attacks on AI-driven systems could disrupt resource allocation, budget management and other critical operations.
- **Public participation**: Where parliaments implement AI to manage large-scale public input, these systems can become targets for manipulation, potentially distorting the representation of public opinion.

- **Data protection**: AI systems often require vast amounts of data, including sensitive information. Ensuring the security and privacy of this data is crucial.
- **AI model integrity**: Attacks could target the AI models themselves, potentially introducing biases or altering decision-making processes without detection.
- **Disinformation campaigns**: AI systems used for information analysis and dissemination could be exploited to spread disinformation within parliamentary networks or to bias research.
- **Autonomous system vulnerabilities**: As parliaments adopt more autonomous AI systems, ensuring that such systems cannot be hijacked or misused becomes critical.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Security management: Threats

About this sub-guideline

This sub-guideline is part of the guideline Security management. It should be read in conjunction with the sub-guideline Security management: Good practices. Refer to the main guideline for context and an overview.

Background

AI systems learn from the data they are fed and then apply models to help them make decisions, generate new content or do anything else they are programmed to do.

For this reason, it is essential that they are fed with correct, clean, unbiased data (for further discussion of this subject, see the guideline Generic risks and biases). Any change to that data, whether intentional or not, may lead to unexpected consequences, results (e.g. a budget management system) or behaviour (e.g. an autonomous car). The end result is akin to teaching improper behaviour or giving wrong information to a child throughout their life.

It is also important to pay attention to the system that will receive user input, send this input to the AI system and then return the result to users. In some cases, attackers can exploit this chain of communication.

Attacks can occur in any phase, from data preparation through to AI system development, deployment and operation (for further discussion of this subject, see the guideline Systems development). As a result, the entire AI system life cycle should be properly supervised in order to minimize unexpected behaviours.

Types of attacks

There are nine common types of attacks on AI systems:

- Adversarial attacks
- Evasion attacks
- Transfer attacks
- Data poisoning attacks

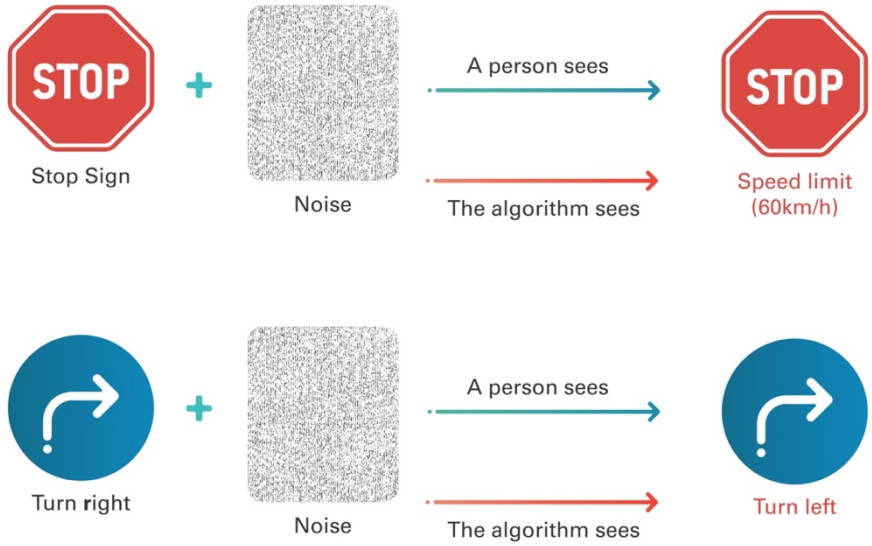
- Model inversion attacks
- Membership inference attacks
- Distributed denial of service (DDoS) attacks
- Data manipulation attacks
- Misuse of AI assistants attacks

Each of these is discussed in turn below.

Adversarial attacks

An adversarial attack involves an attacker manipulating the input data of an AI system so that it produces inaccurate, unexpected or wrong responses. The more information the attacker has about the AI system (especially about the AI model being used), the easier the attack will be. This type of attack often targets AI image recognition systems, making the system incorrectly recognize an image – such as an image of a dog being recognized as a tiger or, worse still, a person being recognized as an animal. The subtle changes in the image are not easily recognizable to the human eye, which makes the issue even harder to solve in certain circumstances. From a parliamentary perspective, an attacker could target a voting system that uses facial recognition technology, causing it to incorrectly allow the attacker to vote as an MP.

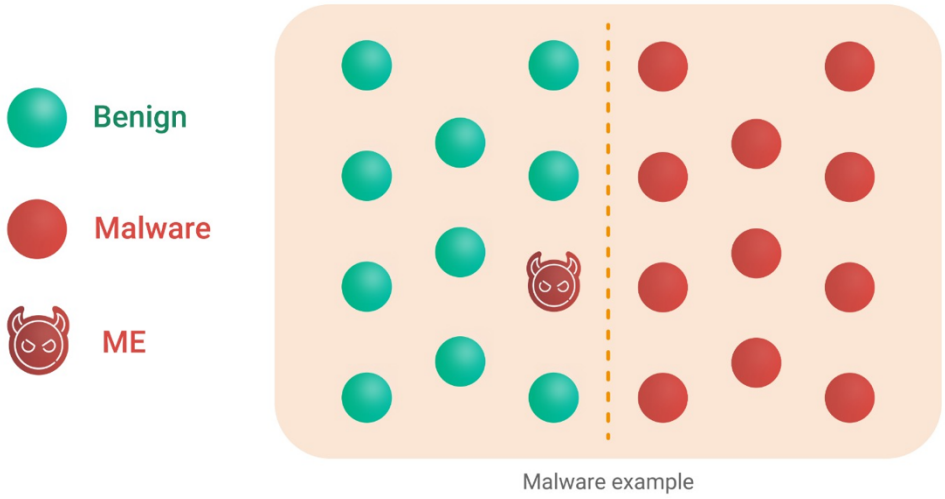
Figure 1: Examples of adversarial attacks



Evasion attacks

An evasion attack is considered to be a specific type of adversarial attack. In this case, the attacker intentionally crafts the input data to evade AI detection or classification. For example, the attacker may change the way an unsolicited email is written to avoid being detected by the AI anti-spam system. This can lead to a malicious message getting through to a regular user, who might click a fake link and allow an attacker to gain access to an organization's network.

Figure 2: Diagram of an evasion attack



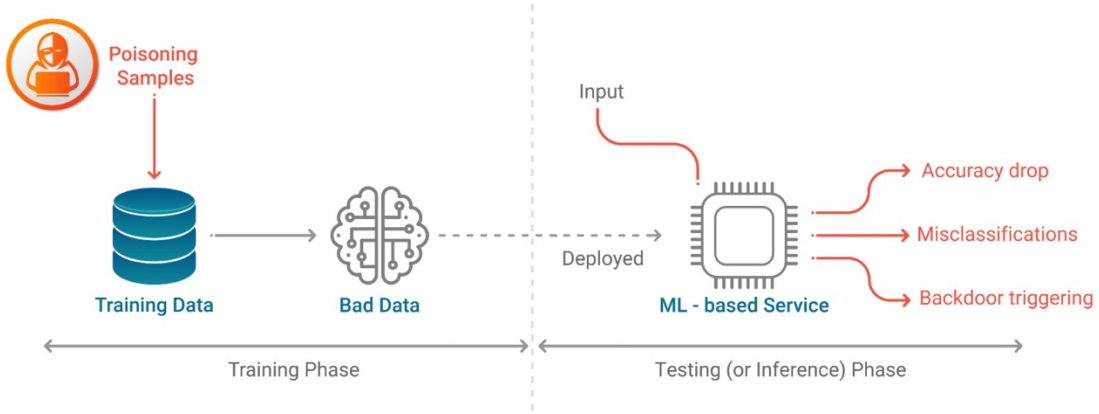
Transfer attacks

A transfer attack occurs when an attacker uses adversarial-type attacks developed for one model and to deceive other models. The consequences are the same as for an adversarial attack.

Data poisoning attacks

In a data poisoning attack, the attacker adds data to the data set used to train an AI model. The model will learn from incorrect information, leading it to make wrong decisions. For instance, a system could wrongly diagnose a healthy patient as having a deadly cancer – or, worse still, wrongly diagnose a patient with cancer as being healthy, preventing the person from receiving proper treatment. In a parliamentary context, a proposal could be forwarded to the wrong committee for discussion.

Figure 3: Diagram of a data poisoning attack

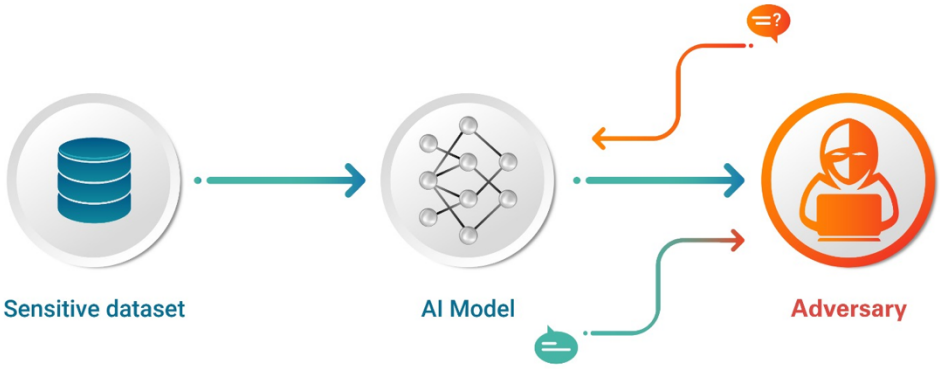


Model inversion attacks

In normal circumstances, an AI model will learn from the input data and produce an output on this basis. The aim of a model inversion attack is to use the output as a way to infer the input data. By doing so, the attacker may gain access to confidential or private information that was used to train the model. For instance, the attacker

could get the result of a specific patient's blood test, or access other, more sensitive data. In a parliamentary context, if an AI model is trained on secret voting data, an attacker may be able to obtain information about how an MP voted.

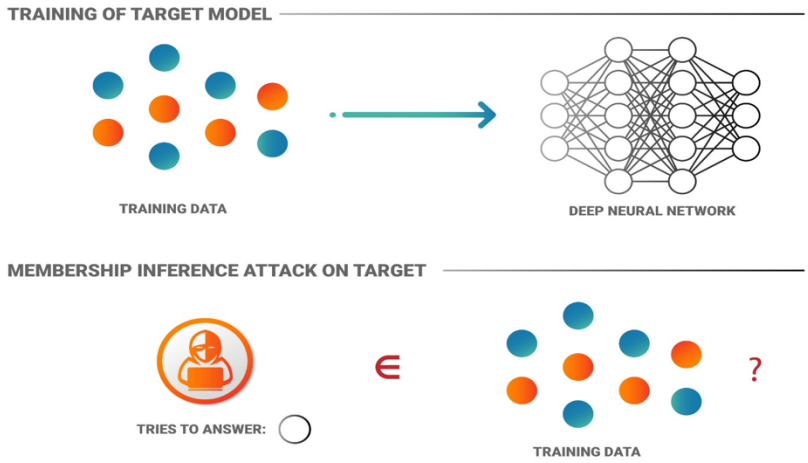
Figure 4: Diagram of a model inversion attack



Membership inference attacks

With a membership inference attack, an attacker aims to find out if individual data records were used to train the AI model. As with a model inversion attack, the attacker may be able to infer sensitive information. For instance, an AI system trained on financial information could leak an individual's financial history.

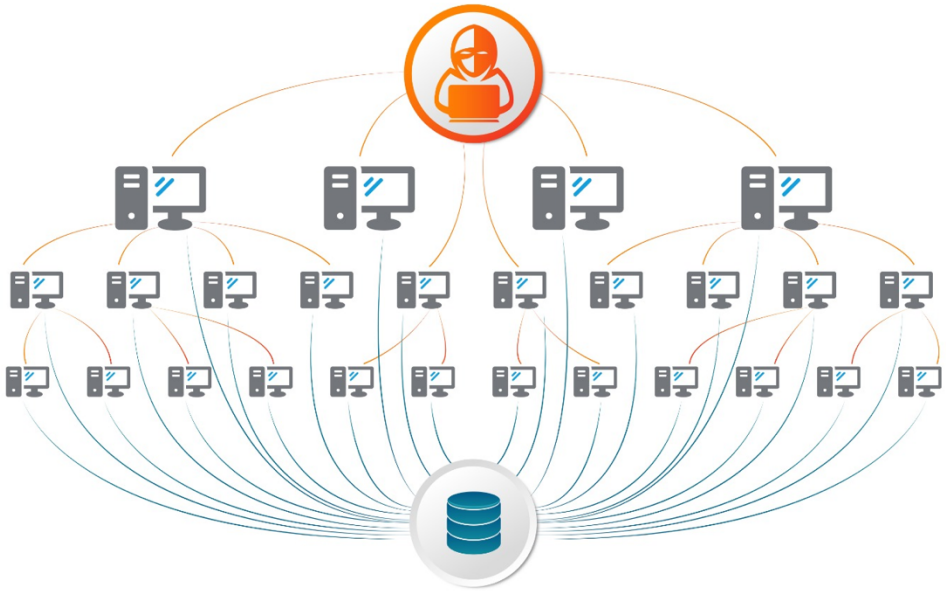
Figure 5: Diagram of a membership inference attack



Distributed denial of service attacks

A DDoS attack involves an attacker flooding a system – including an AI system – with an excessive number of requests. The aim is to cause the system to stop working, preventing any response or, at least, making it so slow that users will not engage. This often leads to financial losses and reputational harm for the organization running the service. In a parliamentary context, an attacker could bring down the AI chatbot designed to answer questions from citizens during a plenary session discussing a theme with broad public support.

Figure 6: Diagram of a DDoS attack



Data manipulation attacks

In a data manipulation attack, the attacker changes the input data (often slightly) in an attempt to generate inaccurate predictions. For instance, an AI system that would normally easily identify a case of fraud may deem such a case to be a regular transaction if the attacker makes minor changes to the input data.

Misuse of AI assistants attacks

The use of AI assistants (such as chatbots or applications built into everyday communication devices) is increasing as these systems become more advanced. It is therefore essential to ensure that a well-thought-out process is in place to select and sanitize the training data set, avoid bias, select the right model for the application, deal with security issues during development, and monitor the application's usage.

In a parliamentary context, a clerk might use an AI assistant (such as the one pre-installed on their mobile phone) to assist them with daily tasks. However, the results this AI assistant produces may be biased according to the clerk's political or other personal preferences.

The problem may become more serious if the AI assistant is attacked by a group with particular preferences on any matter parliament may be discussing, especially if this matter is sensitive. Likewise, if parliament decides to develop an AI assistant to support citizens on legislative matters, it must be considered a target for cyberattacks – not least because its audience is often unknown. In this case, prompts need to be at least sanitized prior to their submission as an input to the AI model, in the same way as inputs to any other AI-enabled system.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Security management: Good practices

About this sub-guideline

This sub-guideline is part of the guideline Security management. It should be read in conjunction with the sub-guideline Security management: Threats. Refer to the main guideline for context and an overview.

Good practices for countering threats to AI systems

Most types of attacks can be avoided or minimized by implementing good security practices. Nonetheless, some attacks targeting AI systems require specific measures. Recommended countermeasures are given below for the following types of attacks:

- Adversarial attacks
- Evasion attacks
- Transfer attacks
- Data poisoning attacks
- Model inversion attacks
- Membership inference attacks
- Distributed denial of service (DDoS) attacks
- Data manipulation attacks
- Misuse of AI assistants attacks

Adversarial attacks

Countermeasures for adversarial attacks, especially those targeting image recognition systems, include the following:

- Use trickier examples in the training phase. For instance, show the AI model lots of slightly altered images so that it learns not to be fooled by them.
- Add a little randomness (technically known as “noise”) to the images used in the training data set. That way, the model will learn to focus on the important parts of the image, not just on the small details that can be easily changed.
- Use stronger models: design the model so that it looks at the big picture (like a person’s overall shape) rather than just focusing on small details.

Evasion attacks

Countermeasures for evasion attacks include the following:

- Choose strong models that are less likely to be fooled by slightly altered inputs.
- Check or validate the input data system to ensure that it is clean and as expected. This can help to catch any abnormal or malicious inputs before they cause harm.
- Train the model using examples of these tricky inputs so that it learns to recognize and handle them properly.
- Regularly monitor how the model performs and update it to handle new types of attacks as they are discovered.
- If possible, and provided that the benefits outweigh the increased cost, use multiple models in combination so that if one model is fooled, the other models can still catch the problem.

Transfer attacks

Countermeasures for transfer attacks include the following:

- Train the model on a wide variety of data. This reduces the chances that an attack crafted on another model will work on the models that parliament is using.
- As with evasion attacks, use multiple models to make decisions – provided that the benefits outweigh the increased cost.
- During training, expose the model to adversarial examples (small, intentionally crafted changes in input designed to fool the model). This helps the model learn to recognize and defend against such attacks.
- If feasible, frequently update and retrain the model with new data. This can help to close any vulnerabilities that might be exploited in transfer attacks.
- Use techniques designed to make the model more resistant to attacks, such as smoothing or noise injection during training.

Data poisoning attacks

Countermeasures for data poisoning attacks include the following:

- Take care over who has access (physical or logical) to the training data set, enforcing robust user permissions.
- Carefully check the data before using it to train the model, including ensuring that the labels and data make sense. Proper sanitization is important for getting rid of data that may negatively impact the learning process.
- Evaluate the machine learning algorithms and check if they are designed to be less sensitive to corrupted data.
- Monitor the model's performance after deployment in order to detect any unusual behaviour that might suggest it was trained on poisoned data.

Model inversion attacks

Countermeasures for model inversion attacks include the following:

- Differential privacy: This technique adds a small amount of random noise to the data or to the model's outputs. The noise is carefully calibrated so that it does not significantly affect the model's performance but makes it much harder for an attacker to extract precise information about individual data points.
- Data minimization: Only collect and use the data that is absolutely necessary for the model. The less sensitive data that is included in the training set, the less risk there is of exposing private information.
- Regularization: This technique can make the model less sensitive to specific data points, which reduces the risk of a successful inversion attack. It forces the model to generalize better, making it harder for an attacker to reconstruct specific inputs.
- Limitation of model access: Restrict who can query the model and how many queries they can make. If an attacker can only make a limited number of queries, it becomes more challenging for them to gather enough information to perform a model inversion attack. A robust account management system plays a key role in defending against this type of attack.
- Query auditing and anomaly detection: Monitor the queries made to the model and look for unusual patterns that might indicate an attack. If suspicious activity is detected, further queries from that source can be blocked.
- Adversarial training: Train the model with adversarial examples (inputs designed to trick the model) to make it more robust against various types of attacks, including model inversion attacks.

Membership inference attacks

Countermeasures for membership inference attacks include the following:

- Regularization: This approach makes the model less confident in its predictions, which makes it harder for an attacker to tell if a specific data point was included in the training data set.
- Differential privacy: This method involves adding noise to the data or to the model's predictions to make it difficult for an attacker to distinguish between data that was included in the training set and data that was not.
- Model distillation: This method trains a simpler model to mimic the behaviour of a more complex model. The simpler model is less likely to give away specific information about the training data.

Distributed denial of service attacks

Countermeasures for DDoS attacks include the following:

- Use a content delivery network (CDN) to distribute the service across servers in different locations.
- Install a web application firewall to detect and block malicious traffic, including a DDoS attack, before it reaches the servers running the AI system.
- Increase server capacity. While this will not solve the problem itself, it will make it more difficult for the attacker to crash the entire system.
- Use externally sourced DDoS protection services to detect and mitigate DDoS attacks. These services can automatically identify and block malicious traffic, keeping a system running smoothly.

- Limit the number of requests a single user can make within a given time frame.

Data manipulation attacks

Countermeasures for data manipulation attacks include the following:

- Enforce the use of strong passwords and multi-factor authentication (MFA) to minimize the risk of unauthorized access.
- Regularly update all software (after proper testing in a controlled environment, as updates can sometimes introduce more problems into the system).
- Keep checking for the most recently discovered vulnerabilities and be ready to fight against them.
- Encrypt all data, and especially data that can lead to privacy issues. Thus, even if attackers are able to access the data, they will not be able to read it without the decryption key (or, at least, it will take a very long time before they can read it).
- Always back up all data properly to avoid data hijacking (which can make the service suddenly stop, negatively impacting parliament's image among citizens and requiring the payment of a high ransom).
- Limit access to sensitive data, and monitor for unusual and suspicious activity such as changes to data or unauthorized logins.

Misuse of AI assistants attacks

Countermeasures for misuse of AI assistants attacks include the following:

- Educate staff to exercise caution as to the information they provide to both personal and enterprise AI assistants.
- Where possible, avoid using AI assistants for AI project purposes.
- If this is not possible, check the privacy settings, as there may be an option to limit the data the AI assistant can access and store.
- Always review the AI assistant's activity logs (if available) for any unusual behaviour.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Security management: Implementing cybersecurity controls

About this sub-guideline

This sub-guideline is part of the guideline Security management. Refer to the main guideline for context and an overview.

Types of cybersecurity controls

Based on the main security frameworks, parliaments should gradually implement controls in the following four areas, according to their specific structure, needs and threat risks:

- Technical controls
- Organizational controls
- Human controls
- Physical controls

Together, measures across these four areas – which are discussed in turn below – enable parliaments to enhance the protection of their AI systems.

Technical controls

Technical controls are measures and processes designed to protect AI systems, data and algorithms from unauthorized access, tampering and exploitation.

Network security

- Use firewalls to segment the network into different zones based on security requirements, implementing strict access controls between zones.
- Consider deploying intrusion prevention systems (IPS) to detect and block malicious activities in real time as soon as possible.
- Ensure that all communication channels – including data transfers, model updates and application programming interface (API) calls – are encrypted in order to protect data in transit.

- Use robust encryption protocols such as transport layer security (TLS) and secure sockets layer (SSL).
- Use virtual private networks (VPNs) to secure remote access to AI systems and data.

System security

- Regularly update and patch all software, operating systems and AI algorithms to protect against known vulnerabilities.
- Be careful to test patches in a controlled environment before deploying them to production systems to ensure they do not introduce new vulnerabilities or cause system instability.
- Deploy robust and regularly updated antivirus and anti-malware solutions on all endpoints, including servers, workstations and mobile devices.
- Enable real-time protection features to detect and block malware and other threats as they occur.

Data security

- Make sure training data sets are reliable and keep these data sets secure, as they are one of the most important assets of the AI system.
- Use data from reliable and verified sources to ensure the authenticity and accuracy of the information.
- If using third-party data, ensure that the data provider has undergone rigorous security audits.
- Remove personally identifiable information from data sets to ensure privacy.
- If this is not possible, replace sensitive data with pseudonyms that can be traced back to the original data only through secure means.
- Encrypt data stored in databases, and in cloud-storage and backup systems, using strong encryption algorithms.
- Use encryption protocols such as TLS or SSL to protect data when it is transmitted between systems or users.
- Pre-process the data to apply sanitization using a variety of methods such as data anonymization, pseudonymization and data masking (for further discussion of this subject, see the guideline Data management).
- Where necessary, establish data-sharing agreements and protocols with trusted partners (such as other parliaments) to ensure the integrity and security of shared data sets, and use secure communication channels when sharing data or collaborating with external parties.

Application security

- Implement systems development best practices to resolve known vulnerabilities and be ready for unknown ones (for further discussion of this subject, see the guideline Systems development).

Organizational controls

Organizational controls focus on internal policies, procedures and practices.

Security policy development

- Develop and implement security policies covering data protection, user behaviour, system access and incident response (for further discussion of this subject, see the guideline Systems development).

Security risk management

- Assess the inherent risks involved in all projects in order to maximize the chances of success (for further discussion of this subject, see the guideline Risk management).

Incident response

- Establish well-defined procedures at a time when the system is not under any real threat. That way, the team can think, discuss and come up with a response plan that is not rushed by the imminent danger.

Human controls

Humans are one of the weakest links in the chain of an AI system, or indeed of any system. Human controls focus on managing this risk through a range of different measures and procedures.

Training and awareness

- Provide security training to the AI team, and ensure that security practices are applied in all stages of the AI system development process within the organization (for further discussion of this subject, refer to the guideline Training for data literacy and AI literacy).

Access management

- Develop a strict role-based access model, implementing the principle of least privilege (PoLP) in order to minimize the risk of unauthorized access and data breaches.

Accountability

- Monitor security incidents and suspicious activities, and implement clear channels for reporting such incidents and activities (for further discussion of this subject, see the guidelines Ethical principles and Systems development).

Physical controls

Physical controls focus on protecting physical assets and infrastructure that support AI systems from unauthorized access, damage or interference.

Facility security

- Ensure that only authorized people have physical access to the AI system.
- Implement at least two ways to allow access the computer room, and apply proper visitor management to sensitive areas.

Environmental controls

- Ensure that facilities hosting IT hardware and staff are protected against fire.
- Install fire detection systems and have an evacuation plan in place.
- If possible, use a climate control system to keep all computers at appropriate temperature and humidity levels.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Risk management

Audience

This guideline is intended for senior parliamentary managers and parliamentary staff involved in the development and implementation of AI-based systems. It also provides a more detailed technical discussion for those involved in developing and implementing AI-related projects.

About this guideline

This guideline provides guidance on AI risk management for parliaments. It emphasizes the importance of continuous risk assessment throughout an AI system's life cycle, from project proposal to decommissioning, with a particular focus on three stages: initial project authorization, the development phase and the operational phase. The associated sub-guideline includes questionnaires to support the identification, analysis and mitigation of AI-related risks.

Relevance of risk management to AI governance in parliaments

There are certain risks associated with the use and development of AI systems in parliamentary settings. Further discussion of these risks can be found in the guidelines Generic risks and biases and Risks and challenges for parliaments.

Managing these risks is essential to ensure that AI systems are safe, ethical, fair, private, trustworthy, transparent and compliant with regulations, as well as to ensure respect for human autonomy and intellectual property rights. In summary, effective AI risk management practices help parliaments to:

- protect their data and AI systems
- maintain business continuity
- safeguard their reputation
- prevent costly errors
- support responsible innovation

Planning to implement an AI risk management process

An AI risk management process needs to be aligned with parliament's culture and structure. As stakeholders are at different hierarchical levels, appropriate language

must be used in assessment questionnaires and periodic reports in order to prevent delays or incorrect interpretations.

During the development phase, adjustments to traditional risk management processes, embedded in project management, can simplify and speed up the implementation of AI risk management. In this case, the process will involve not only AI project managers, but also IT committees, corporate committees and senior decision makers.

During the operational phase, a partnership between business and IT teams facilitates the integration of oversight practices and the collection of user feedback as input for AI risk management. In this case, it is crucial to involve a risk management team composed of people with AI skill and business staff in charge of AI-enabled digital services.

Applying risk management processes to the AI life cycle

Risk management is a continuous cycle that permeates all phases of an AI system's life cycle, ensuring that risks are consistently identified, assessed, and managed or mitigated – from inception to deployment and beyond.

A typical AI risk management process includes the following phases:

- AI risk assessment
- AI risk analysis
- AI risk treatment
- AI risk communication
- AI risk monitoring and review

These phases are discussed in turn below.

AI risk assessment

The aim of this first step is to identify what risks exist, to understand the operational and other implications of these risks, and to decide how they should be managed or mitigated.

This assessment exercise will typically be carried out using questionnaires for key stakeholders – which, in itself, can be a useful mitigation method because it exposes potential risks and increases awareness of them.

These questionnaires can be used in the following phases:

- Initial project authorization
- The development phase (prior to commissioning)
- The operational phase (through to decommissioning)

Figure 1: AI risk assessments throughout the AI system life cycle

Risk assessment for initial project authorization

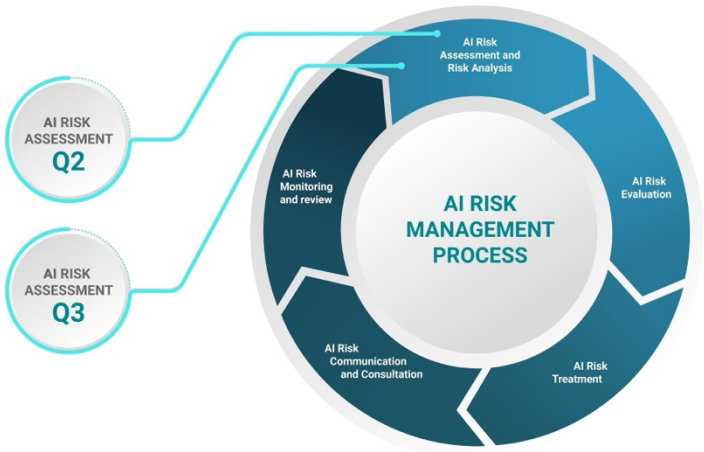
The AI system life cycles starts with the presentation of a proposal for an AI project to the relevant governance body (council, committee or unit), along with a completed AI risk assessment questionnaire (Q1), which is used to gather information about the project's purpose, stakeholders, compliance, data, agreements, potential biases and other factors.

Governance staff use the responses in the questionnaire to estimate the project's risks and benefits, and to determine, on that basis, whether authorization should be given to add the AI project to parliament's portfolio.

Risk assessment in the development phase

During the development phase, the Q2 questionnaire will inform the risk management process, with the aim of reducing and mitigating AI risks such that the output AI system is considered trustworthy for deployment. Unacceptable AI risks can, however, lead to the project being interrupted at this stage.

Figure 2: Q2 and Q3 assessment questionnaires used in the risk management process



Risk assessment in the operational phase

After the AI system has been deployed, it is necessary to keep monitoring the system's behaviour as well as changes in the variables considered in the AI system life cycle, such as data characteristics, business rules and social considerations.

The third risk assessment questionnaire (Q3) can be used during this phase. Similarly to the Q2 questionnaire, the risk score resulting from this third questionnaire will inform the risk management process, which at this point aims to reduce and mitigate AI risks in order to ensure that the system remains trustworthy.

AI risk analysis

Once the risks have been identified and assessed, the next step is to analyse these risks in light of parliament's AI policy and its risk appetite – often based on its regulatory requirements and strategic objectives – in order to determine which risk(s) require(s) treatment. All identified risks should then be ranked in order to identify which require immediate attention and which should be monitored over time. All such decisions should be made with the close involvement of relevant stakeholders.

Trade-offs between different risk treatment options also need to be evaluated at this stage. For instance, by eliminating one identified risk, parliament could be at risk of not achieving other strategic goals by also eliminating the option of using the AI system in another, important way.

AI risk treatment

Once the identified risks have been analysed and prioritized, parliament should develop and implement plans to manage them, possibly using one or more of the following strategies:

- **Avoid**: Eliminate the activity that gives rise to the risk. For example, parliament may decide not to implement an AI system, or even to abandon an AI project, if the associated risks are deemed too high.
- **Reduce**: Take steps to reduce the likelihood of the risk occurring, or to mitigate its impact if it does occur.
- **Transfer**: Transfer the risk to a third party, such as through insurance or by outsourcing certain services to a company better equipped to manage the risk.
- **Accept**: Accept the risk without taking any action to alter its likelihood or impact. This is typically done when the cost of mitigating the risk exceeds the potential damage, and when the risk is considered low enough to be acceptable.

Figure 3: Risk treatment strategies

AI risk communication

Identified AI risks and associated management measures should be communicated to relevant stakeholders throughout the AI system's life cycle. During the development phase, project managers and project office staff will provide regular updates on risk status and treatment effectiveness as part of their usual remit. Communication is equally important in the operational phase.

AI risk monitoring and review

During the operational phase, it is essential to continuously monitor and review AI risks.

Oversight and feedback mechanisms, coupled with training programmes, help to build a risk-aware culture and keep stakeholders informed.

Periodic audits and reviews should also be conducted to ensure compliance with AI policy and regulations. All identified incidents and near-misses should be analysed in order to identify root causes and improve risk management practices, with lessons learned documented and policies updated accordingly.

Should any unacceptable AI risks arise, it may be necessary to remove the AI system from operation.

Find out more

- Cheatham, B., Javanmardian, K., and Saman, H. (Undated). Confronting the risks of artificial intelligence, available at [Confronting AI risks | McKinsey]
- National Institute of Standards and Technology (NIST) Special Publication (SP) 800-30, Guide for Conducting Risk Assessments, available at [NIST SP 800-30 | NIST]
- Online Browsing Platform (OBP), ISO/IEC 27005:2022(en) Information security, cybersecurity and privacy protection — Guidance on managing information security risks, available at [ISO/IEC 27005:2022(en), Information security, cybersecurity and privacy protection — Guidance on managing information security risks]
- Tucker, B.A. (undated). Carnegie Mellon University, Advancing Risk Management Capability Using the OCTAVE FORTE Proces, available at [Advancing Risk Management Capability Using the OCTAVE FORTE Process (cmu.edu)]

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Risk management: Risk assessment questionnaires

About this sub-guideline

This sub-guideline is a part of the guideline Risk management. Refer to the main guideline for context and an overview.

This sub-guideline provides sample risk assessment questionnaires that can be used to support parliament's risk management process in three phases of the AI system life cycle:

- Initial project authorization
- The development phase
- The operational phase

Question 1: AI risk assessment for initial project authorization

When business managers want to submit a proposal to run an AI project, through a partnership with the IT unit, they focus on the problem to be solved using AI. At this stage, both business managers and the IT unit should complete an initial questionnaire, which could include the following questions, among others:

Purpose and stakeholders

- What is the business case for, and the problem to be solved by, the proposed AI system?
- Which stakeholders would this project benefit? How would they be impacted?
- Are there any stakeholders (internal and/or external) that could be negatively impacted by this project? If so, could these negative effects be mitigated or compensated for?

Compliance

- Will this project conflict with parliament's policies, or with any law or other compliance rule?

Data and privacy

- Have the owners of the data that the AI system will use been identified?
- Will the AI system use internal data? If so, has internal authorization been obtained to use parliament's data for this purpose?
- Will the AI system use external data? Has parliament signed a memorandum of understanding or other agreement with the organization(s) that own(s) the data?
- Are any groups potentially underrepresented in the data?
- Will the AI system use personal data? Is there an agreement or arrangement with appropriate safeguards in place?
- Will the AI system's output be available to external users?

Copyright

- Should any of parliament's data be protected by copyright?
- Are there any copyrights or contractual conditions that need to be respected?

Capacity-building and outsourcing

- What expertise is missing within parliament (if any) to support the procurement, development or implementation of the AI system?
- Will this AI system be developed internally, purchased as a commercial product or developed through outsourcing?

Question 2: AI risk assessment in the development phase

This second questionnaire is to be completed by the AI development team and business staff during the development phase. Where an agile development method is employed, the responses should be reviewed for each new version of the system, informing the decision as to whether to continue or suspend – or even cancel – the project. This questionnaire could include the following questions, among others:

Data and privacy

- Have the data owners authorized all actions regarding data access and treatment?
- Is enough data available for the project?
- Is the data quality adequate for this project? If not, what data quality issues have been identified?
- How will parliament improve data quality in relation to the identified issues?
- Will the use of personal data be restricted to the purposes for which it was planned/authorized?
- Is it possible to keep people's privacy sufficiently protected? Will it be possible to re-identify the data subjects?

Bias and discrimination

- Are there any underrepresented data categories or potential biases in the data set? If so, what issues have been detected and how are they being mitigated?
- If generative AI is used, which hallucinations should be avoided or minimized in this project?
- Will the AI system generate any classification of people's behaviour?

Transparency

- How are the planning, modelling, evaluation, testing and deployment phases scrutinized?
- Does the documentation use appropriate language for the target audience(s)?
- How do the documents demonstrate that the model addresses business requirements?
- How do the documents demonstrate that the AI system is sufficiently accurate?
- Is there any direct interaction between human end users and the AI system? Are users explicitly informed that they are interacting with an AI system?

Safety and robustness

- Are there any weaknesses in the defined model?
- Are there any weaknesses in the testing phase?
- Is the AI system robust to potential failures and security attacks?
- Is there a deployment plan?
- Is there a rollback or disaster recovery strategy in place?

Question 3: AI risk assessment in the operational phase

Once an AI system is deployed in a live operating environment, it should be continuously monitored by business and IT staff. At this stage, these staff should complete a third questionnaire to ensure that changes in data, business rules, social trends and the operating environment have been taken into account. This questionnaire could include the following questions, among others:

Human autonomy and oversight

- Is the AI system subject to continuous performance monitoring?
- Has parliament established clear criteria for classifying acceptable and/or unacceptable AI system behaviours? If so, is the AI system's current behaviour acceptable according to these criteria?
- Has parliament implemented a continuous process for collecting user feedback on the AI system's behaviour? If so, is the AI system's current behaviour acceptable according to user feedback?
- Has parliament identified any new variables (changes) in the AI system that were not considered in the development phase?

Transparency

- Is the performance monitoring and user feedback collection process effectively scrutinized?

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Systems development

Audience

This guideline is intended for IT managers and staff, software engineers, data scientists and technical project managers involved in designing, deploying and maintaining AI systems in parliaments.

About this guideline

This guideline focuses on AI systems development within parliamentary contexts, addressing the crucial intersection of technology and governance. It covers essential aspects such as the AI system life cycle, external development frameworks, deployment strategies and planning considerations. By emphasizing ethical principles, risk mitigation and best practices, this guideline aims to support parliaments in implementing AI solutions that enhance efficiency, transparency and decision-making while maintaining integrity and public trust in the legislative process.

Why AI systems development matters

In the context of AI governance, an AI systems development process is a set of practices designed to ensure that all AI projects solve the problems for which they were planned and adhere to AI ethical principles. As an inherently operational process, AI systems development should adhere to parliament's AI policy, as well as follow the institution's data management and security management procedures.

In a parliamentary context, an AI systems development process is relevant to AI governance as a way of reducing ethical and operational risks. Details of how it does this are given below.

Preserving the privacy of data subjects

Personal data should be protected not only in the development phases, but also in the AI system's outputs.

Ensuring transparency throughout the development and maintenance phases

AI systems often imply complexity, making their internal decision-making processes opaque. For this reason, practices should be scrutinized in order improve transparency throughout the AI system life cycle – from initial project planning until the AI system is withdrawn from operation. This approach should make it easier to

explain the AI system's outcomes and ensure that the development phases respect parliament's rules and are compliant with regulations.

Reducing biases and discrimination

Techniques to identify groups that are to be protected from biases should be applied throughout the process, from planning to deployment. While AI systems are in operation, continuous monitoring helps to minimize new biases not seen in the development phase.

Creating accountability

Through systematic steps for planning, implementing, testing and improving, practices are delegated and approved by key stakeholders, with individual roles and responsibilities clearly defined. The AI system's functionality should be documented in such a way that it can be audited.

Improving robustness and safety

The systems development process should focus on improving robustness and safety, through a system architecture that prioritizes cybersecurity and through extensive testing.

Maintaining human autonomy

Humans should play a continuous verification role in order to ensure that the AI system's outputs are reliable, both during development and following deployment in a live environment. This human oversight will ensure that the system continues to adhere to the ethical principles considered in the project phase, and allows for new ethical risks to be identified.

Guaranteeing regulatory compliance

AI systems are required to comply with various legal and regulatory requirements, established both internally and within parliament's country or region.

Systems life cycle and development frameworks

The AI systems life cycle is a sequential list of steps, practices and decisions that drive the development and deployment of AI-based solutions. Having a well-defined life cycle is vital for parliaments that are developing their own AI-based systems and tools, as it provides a structured and systematic approach to building, deploying and maintaining ethical AI technologies.

Within this context, there are an increasing number of external AI development frameworks that parliaments can use. These consist of building blocks and integrated software libraries that make it easier to develop, train, validate and deploy AI solutions through a high-level programming interface.

For further discussion of systems life cycle and development frameworks, refer to the sub-guideline Systems development: Systems life cycle and development frameworks.

Deployment and implementation

When deploying and implementing AI systems and tools, parliaments need to understand key aspects such as deployment strategies, common deployment cases and critical planning recommendations, including topics such as stakeholder engagement, pilot project initiation and the use of agile methods. Context should also be given consideration, including parliamentary workflows, internal expertise and opportunities for leveraging responsible AI tools.

For further discussion of the deployment and implementation of AI systems, refer to the sub-guideline Systems development: Deployment and implementation.

For further discussion of software deployment patterns, refer to the sub-guideline Systems development: Deployment patterns.

Find out more

- Data Science PM: "The GenAI Life Cycle"
- Data Science PM: "What is CRISP DM?"
- Data Science PM: "What is the AI Life Cycle?"
- Google: "Domain-specific AI apps: A three-step design pattern for specializing LLMs"
- Inter-American Development Bank (IDB): *Responsible use of AI for public policy: Data science toolkit*
- Microsoft: "Empowering responsible AI practices"
- Microsoft: "What is the Team Data Science Process?"
- Wikipedia: "Cross-industry standard process for data mining"

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Systems development: Systems life cycle and development frameworks

About this sub-guideline

This sub-guideline is part of the guideline Systems development. Refer to the main guideline for context and an overview.

This sub-guideline focuses on the systems life cycle and development frameworks for AI-based systems in parliamentary contexts. It provides an overview of the AI systems life cycle, highlighting its importance in ensuring structured and responsible AI development.

This sub-guideline outlines the benefits of adopting a systematic life cycle approach. It also offers guidance on evaluating external AI development frameworks, covering aspects such as ease of use, community support, performance, model support and deployment readiness.

This sub-guideline is intended to help parliamentary IT professionals make informed decisions about AI development processes and tool selection, ultimately supporting the effective and responsible implementation of AI technologies in legislative environments.

The AI systems life cycle

The life cycle of an AI system is a sequential list of steps, practices and decisions that drive the development of AI-based solutions. Having a well-defined life cycle is vital for parliaments that are developing their own AI-based systems and tools, as it provides a structured and systematic approach to building, deploying and maintaining ethical AI technologies.

Specifically, adopting an AI systems life cycle approach offers the following benefits:

- Increased success rate: Following each essential step in the development of an AI system improves the chances of project success.

- Risk reduction: Identifying potential issues early in the process helps to mitigate risks and prevent costly setbacks.
- Improved efficiency and productivity: An organized project timeline makes work smoother, ensuring that all team members understand their roles and responsibilities at each stage.
- Enhanced quality: Completeness and rigour at each stage of the life cycle lead to higher-quality AI systems.
- Better resource allocation: AI projects require significant resources, including time, human expertise and computational power. Properly identifying and balancing these resources ensures that they are used effectively throughout the project.

External development frameworks

There are an increasing number of external AI development frameworks that parliaments can use. These consist of building blocks and integrated software libraries that make it easier to develop, train, validate and deploy AI solutions through a high-level programming interface.

The natively pre-configured blocks and functions provided by these frameworks speed up implementation time. By allowing developers to solve tasks by customizing existing blocks without having to start from scratch, these frameworks also improve productivity and algorithm quality. Moreover, using standard frameworks makes it easier to integrate AI features with a great variety of application platforms and domains.

In order to compare and evaluate different external AI development frameworks, parliament must understand their characteristics and determine their suitability for its workflows and business needs. It is advisable to review specific frameworks for specific use cases, and to compare options through experimentation.

Key considerations for this decision-making process are detailed below.

Ease of use

- Documentation: Quality, clarity and comprehensiveness of the documentation
- Learning curve: How easy it is to start using the framework, including the availability of tutorials and community support
- API design: Simplicity and intuitiveness of the API

Community and support

- Community size: The number of users and developers contributing to the framework
- Support: The availability of forums, user groups and other support channels
- Updates: The frequency of updates, how many known issues and vulnerabilities exist, and how actively the framework is maintained and improved

Performance

- Speed: How quickly models can be updated for training and inference
- Scalability: The ability to manage large data sets and complex models, and support for distributed training
- Optimization: Built-in features for optimizing and tuning model performance and resource usage

Model support

- Model variety: The range of supported model types (neural networks, decision trees, etc.)
- Pre-trained models: The availability and variety of pre-trained models that can be fine-tuned or used out of the box
- Customization: Flexibility in defining and experimenting with custom models and architectures

Tooling and integration

- Ecosystem: The availability of complementary tools for data preprocessing, visualization and deployment
- Compatibility: Integration with data-handling libraries, visualization tools, deployment platforms, etc.
- Interoperability: Support for importing/exporting models between different frameworks

Deployment and production readiness

- Deployment options: Ease of deploying models to different environments (cloud, edge, mobile)
- Serving: Support for model serving and inference in production settings
- Monitoring: Tools for monitoring model performance and detecting issues in production
- Data protection regulations: Assurance that data classification, retention and residency rules are followed

Licensing and cost

- Open-source versus proprietary: Whether the framework is open-source or commercial
- Licensing terms: Any restrictions or requirements imposed by the licence
- Cost: The potential costs associated with using the framework, especially for proprietary options

Hardware support

- GPU/TPU support: Compatibility with various hardware accelerators
- Distributed computing: Support for running on multiple GPUs or across a cluster of machines

Extensibility

- Plugins: The availability of plugins or extensions for added functionality
- APIs for custom extensions: The ability to write custom extensions or integrate third-party tools

Scalability

- Scaling up and out: Support for horizontal and vertical scaling (manual or automatic)
- Performance: Load testing and simulations to measure the performance of the framework
- Cost: The ability to set costing limits in the event that the system needs to scale

Reproducibility

- Versioning: Tools for model and data versioning to ensure the reproducibility of outcomes
- Experiment management: Support for tracking experiments and managing their results

Security

- Security features: Built-in security features for safe deployment and model usage
- Compliance: Compliance with industry standards and regulations

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Systems development: Deployment and implementation

About this sub-guideline

This sub-guideline is part of the guideline Systems development. It can be read in conjunction with the sub-guideline Systems development: Deployment patterns. Refer to the main guideline for context and an overview.

This sub-guideline focuses on the deployment and implementation of AI systems and tools within parliament. It provides essential insights for IT professionals and decision makers involved in integrating AI solutions into parliamentary operations.

This sub-guideline covers key aspects such as deployment strategies, common deployment cases and critical planning recommendations. By addressing topics like stakeholder engagement, pilot project initiation and the use of agile methods, this guideline aims to support parliaments in effectively and responsibly implementing AI technologies to enhance legislative processes, improve efficiency and maintain transparency. It emphasizes the importance of understanding parliamentary workflows, building internal expertise and leveraging responsible AI tools throughout the implementation process.

Deployment strategy

A structured and coordinated approach to AI systems deployment is essential to the effective integration of these systems into various parliamentary applications and workflows. These patterns of deployment reinforce good practices, helping to ensure that AI deployments are scalable, robust and maintainable. For further discussion of software deployment patterns, refer to the sub-guideline Systems development: Deployment patterns.

Parliament's deployment strategy will depend on the degree of task automation. The various options are discussed below:

- Human-only: In this case, there is no automation. The task is carried out manually by users without AI support.

- AI assistance: The task is performed mainly by users, possibly with assistance and support from the AI system.
- Partial automation: The task is performed mainly by the AI system, which produces suggestions for users.
- Full automation: The task is performed entirely by the AI system, without human intervention.

Both AI assistance and partial automation are examples of “human-in-the-loop” deployments (for further discussion of this topic, refer to the sub-guideline Ethical principles: Human autonomy and oversight).

Deployment cases

The most common deployment cases and their characteristics are detailed below:

- New product or feature: In this case, a new AI-based product or feature is introduced.
- Partial task automation: In this case, a task was previously done manually and an AI algorithm is introduced to either automate this task or assist the user.
- Replacement of a previous AI system: In this case, a task was carried out via a previous implementation of an AI system, and another AI system is introduced to replace the previous one with a view to improving quality and/or execution time.

In the above deployment cases, parliament should consider the following two basic aspects:

- Gradually increase traffic with monitoring. It is advisable to avoid sending a lot of production traffic to an algorithm that is still learning and is not yet fully proven. It may be better to send it only a small amount of traffic, monitor it and then progressively ramp up the amount of traffic.
- Ensure there is a possibility to roll back. It is advisable to have a contingency plan in place to revert back to the previous, stable configuration in case the new algorithm does not work as expected.

Planning the implementation of AI systems

Key recommendations for planning the implementation of AI systems within a parliamentary context are outlined below.

Understand parliamentary processes

- Conduct a comprehensive analysis of existing parliamentary workflows and challenges.
- Examine key processes such as legislative drafting, committee meetings, voting procedures and public consultations.
- Map out the flow of documents, communications and decisions within parliament.
- Understand how information is processed and identify things such as pain points, redundancies and inefficiencies in current processes.
- Understand the current state of digital systems, databases and communication tools used for parliamentary operations.

- Identify where AI can be integrated to enhance current systems or replace outdated technologies.
- Identify and develop potential use cases to improve understanding of the potential for AI solutions at both technical and business levels.
- Ensure that any AI implementation complies with existing laws and with parliament's rules (or identify where changes to these rules are needed).

Engage stakeholders

- Ensure that the following stakeholders are represented and involved throughout the entire AI system life cycle:
  - Business experts
  - MPs, clerks and administrative staff, who will use the AI system for legislative activities
  - Legal experts, who will ensure that the AI implementation adheres to relevant laws
  - Citizens, who need to be informed and consulted about AI initiatives in parliament
  - External experts such as AI researchers or data scientists, who can provide insights into AI implementation
- Regularly solicit feedback from stakeholders at all stages of the AI project and incorporate this feedback into the AI development process.
- Keep stakeholders informed about the progress of AI projects, milestones achieved, and any changes or updates, through regular reports and presentations.

Start with a pilot project

- Initiate small-scale pilot projects to test AI solutions in real parliamentary settings.
- Choose pilot projects with a clear and manageable scope.
- Look for areas within parliamentary processes that can benefit from AI, such as document analysis, constituent communication or data-driven decision support.
- Assess the feasibility and potential impact of proposed pilot projects.
- Establish clear, specific objectives for each pilot project. These should be aligned with the overall goals of enhancing parliamentary efficiency, transparency and decision-making.
- Develop success criteria to measure the effectiveness of the pilot projects. These criteria could include performance improvements, AI output accuracy, user satisfaction and cost savings.
- Create a detailed project plan that outlines the steps, timelines, resources and responsibilities for the pilot project.
- Identify potential risks and develop mitigation strategies.

Use agile methods

- Plan the activities in sprints, each lasting a couple of weeks.
- Involve business experts in each phase of the project. Doing so offers numerous benefits including a faster response to legislative changes,

better cross-team collaboration, higher user satisfaction, reduced project risks and increased transparency for stakeholders.

Build internal AI expertise

- Invest in developing AI expertise among parliamentary staff.
- Provide training and resources to help staff understand, evaluate and effectively use AI technologies.

Collaborate with AI experts and share with other parliaments

- Collaborate with AI experts, researchers and technology providers to build knowledge and experience.
- Collaborate with academic institutions, research organizations and other parliaments to share and learn, and to access cutting-edge AI research and innovations.

Leverage responsible AI tools

Consider using the recommended tools detailed in the Inter-American Development Bank publication *Responsible use of AI for public policy: Data science toolkit*:

- Robust and Responsible AI Checklist: This tool consolidates the main concerns by stage of the AI life cycle. The checklist must be reviewed continuously by technical teams and decision makers.
- Data Profile: This tool is an initial exploratory analysis conducted during the data-collection and processing stage of the AI life cycle. It provides information to reassess the quality, completeness, temporality and consistency of the training data set, possible biases within this data set, and the implications of the use of an AI system, including potential harm.
- Model Card: This tool summarizes the main features of the AI system, highlighting the main assumptions, the most important characteristics of the system, and the risk mitigation measures implemented.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Systems development: Deployment patterns

About this sub-guideline

This sub-guideline is part of the guideline Systems development. It can be read in conjunction with the sub-guideline Systems development: Deployment and implementation. Refer to the main guideline for context and an overview.

Background

The choice of technology and infrastructure significantly influences the robustness, security, scalability, performance and ease of supervision of AI systems. Together, these factors contribute to the overall reliability of such systems.

Different deployment patterns come with their own set of trade-offs. Some may incur higher costs, while others might require more extensive management resources. As a result, there is no one-size-fits-all "best approach" to AI system deployment. Instead, the key to successful deployment lies in carefully assessing parliament's needs, resources and goals, and then selecting an approach that offers the best balance of features and practicality for parliament's particular situation.

Characteristics of deployment patterns

The deployment of AI systems involves following several patterns and practices to ensure that models perform effectively and reliably in production environments. When designing the AI system architecture, parliament should therefore consider the deployment pattern characteristics discussed below.

Deployment architecture

The deployment architecture of an AI system is determined by two key factors: how the AI algorithm responds to requests, and where the AI model is hosted.

Request-handling patterns:

- Batch processing: Data is processed in large batches at scheduled intervals, making this pattern suitable for non-time-sensitive tasks.
- Online serving: Requests are handled in real time as they come in, making this pattern ideal for applications requiring immediate responses.

- Streaming: Under this pattern, data streams are continuously processed, enabling near-real-time analysis and predictions.

Hosting location types:

- On-premises: Models are deployed on local servers, often for the purpose of enhanced security or to meet specific compliance requirements.
- Cloud: Models are hosted on cloud platforms, offering benefits such as scalability, flexibility and reduced infrastructure management.
- Edge: Models are deployed on edge devices, providing low-latency predictions and offline capabilities, making this approach suitable for Internet of Things (IoT) and mobile applications.
- Hybrid: This approach combines on-premises, cloud and edge deployments to optimize performance and resource usage based on specific needs.

The choice of deployment architecture depends on factors such as data sensitivity, response-time requirements, available resources and the specific use case of the AI system.

Scalability

It is important to understand the average number of requests the AI system will receive, along with its life cycle. These factors will determine the deployment scalability characteristics:

- Horizontal scaling: Adding more instances of the model server to handle increased load
- Vertical scaling: Enhancing the capacity of existing servers (e.g. by adding more memory or faster central processing units (CPUs))
- Auto-scaling: Automatically adjusting the number of model instances based on demand

Latency and throughput

When deploying AI systems, two critical performance metrics to consider are latency and throughput:

- Latency refers to the time it takes for the AI model to respond to a request, which is particularly crucial for real-time applications.
- Throughput measures the number of requests the AI model can process per unit of time, which is essential for high-volume applications.

It is important to establish acceptable values for both latency and throughput to ensure that the system meets the specific needs of the application for which it is intended, and that it can handle the expected workload efficiently.

Model management

Effective AI model management is crucial throughout the entire life cycle of an AI system. However, it becomes particularly important once the AI system is put into

operation. A well-designed model management strategy should address several key aspects:

- Versioning: This involves keeping track of different versions of the model, ensuring traceability and the ability to roll back if needed. Proper versioning allows teams to manage changes, compare performance across iterations and maintain a clear history of the model's changes over time.
- Life cycle management: This approach encompasses the tools and processes for deploying, monitoring, updating and, eventually, retiring models. The aim is to ensure that models are properly maintained throughout their operational life, from initial deployment through to eventual replacement.
- A/B testing: This practice involves running multiple versions of a model simultaneously to compare their performance. A/B testing allows teams to make data-driven decisions about which model version performs best in real-world conditions before full deployment.

Monitoring and observability

- Performance metrics: Monitoring metrics such as response time, throughput and resource utilization
- Drift detection: Identifying when the model's performance degrades owing to changes in data distribution
- Alerting: Setting up alerts for anomalies or performance degradation

Security

- Access control: Ensuring that only authorized users and applications can interact with the model
- Data privacy: Protecting sensitive data and adhering to regulations (e.g. GDPR)
- Model security: Safeguarding models against adversarial attacks and data poisoning

Continuous integration/continuous deployment (CI/CD)

- Automation: Automating the deployment process to reduce errors and deployment time
- Testing: Including automated testing (unit, integration, regression) in the deployment pipeline
- Rollbacks: Providing mechanisms for quickly reverting to previous versions in case of issues

Resource management

- Hardware acceleration: Utilizing graphics processing units (GPUs), tensor processing units (TPUs) or other accelerators for improved performance
- Resource allocation: Managing resources to optimize cost and performance
- Integration with existing systems: Providing APIs for integration with other systems and services

- Data pipelines: Integrating with data ingestion and pre-processing pipelines
- Feedback loops: Implementing systems to collect feedback from model predictions to improve future performance

Resilience and fault tolerance

- Redundancy: Having multiple instances or backups to ensure availability
- Failover: Automatically switching to backup systems in case of failure
- Retry logic: Implementing mechanisms to handle transient failures

Auditability and explainability

In most cases, audit logs are mandatory for predictions, inputs and system interactions.

In addition to auditing, explainability tools can be used to interpret AI model decisions, thus improving trust and compliance.

Combinations of deployment patterns

Various combinations of characteristics are often seen in AI use cases. These are detailed below:

Model-as-a-service (MaaS)

- Characteristics: Exposing models via web APIs for easy integration
- Use cases: Real-time predictions, microservices architecture

Model embedded in applications

- Characteristics: Embedding models directly in applications, either locally or via a microservice
- Use cases: Edge computing, offline capabilities, low-latency requirements

Containerized deployment

- Characteristics: Packaging models in containers (e.g. Docker) for consistent deployment across environments
- Use cases: Cloud deployments, microservices, scalable architectures

Serverless deployment

- Characteristics: Using serverless computing platforms to deploy models
- Use cases: Event-driven applications, cost optimization for intermittent workloads

On-demand/batch processing

- Characteristics: Deploying models that run on demand or process large batches of data periodically
- Use cases: Data-processing pipelines, periodic analytics

Streaming analytics

- Characteristics: Deploying models to analyse and predict data from streaming sources

- Use cases: Real-time analytics, IoT applications

A/B testing and canary releases

- Characteristics: Testing new models on a subset of traffic before full deployment
- Use cases: Incremental updates, risk minimization

Federated learning

- Characteristics: Training models across multiple decentralized devices or servers while keeping data local
- Use cases: Privacy-sensitive applications, distributed data sources

Recommended practices when deploying AI systems

Below are some suggestions and characteristics for parliaments to consider when planning and executing AI deployments that are scalable, reliable, safe and efficient:

- **Ensure best fit**: Select a deployment pattern according to the specific use case, performance requirements and domain constraints.
- **Monitor and iterate**: Continuously monitor deployed models and iterate based on user feedback and performance metrics.
- **Maintain security**: Implement robust security practices to protect models and data in production environments.
- **Optimize resources**: Efficiently manage resources to balance performance and cost, leveraging approaches such as containerization and serverless architectures where appropriate.

Inter-Parliamentary Union
For democracy. For everyone.

Guidelines for AI in parliaments

Appendices

Guidelines for AI in parliaments

Glossary of terms

Accountability	The principle that ensures clear responsibility can be assigned for all decisions and actions throughout an AI system's lifecycle, from planning to decommissioning
Affinity bias	When someone prefers individuals who are similar to them in terms of ideology, attitudes, appearance, or religion
Agile	A project management and development approach that emphasizes flexibility, iterative progress, and collaboration
AI governance	The framework of policies, structures, and processes created to maximize the benefits of AI while minimizing its risks
AI literacy	The ability to understand, critically evaluate, and effectively interact with AI technologies, including knowledge of AI's capabilities, limitations, and potential impacts
AI PPM (AI Project Portfolio Management)	The centralized management of an organization's AI initiatives to meet strategic objectives by optimizing resource allocation, balancing risks, and maximizing value
Algorithm	A set of rules or instructions given to an AI system to help it learn, make decisions, and solve problems
Amplification bias	Occurs when several AI systems, each with separate biases, interact and mutually reinforce each other's biases
API (Application Programming Interface)	A set of rules and protocols that allows different software applications to communicate with each other
Automation bias	When conclusions drawn from algorithms are valued more highly than human analyses
Cloud storage	The practice of storing data and applications on remote servers accessed via the internet, rather than on local computers
Coverage bias	A form of sampling bias that occurs when a selected population does not match the intended population

Data architecture	The overall structure of an organization's data assets and data management resources
Data bias	A type of error where certain elements of a data set are more heavily weighted or represented than others
Data cleaning	The process of detecting and correcting errors, inconsistencies, and inaccuracies in data sets
Data governance	The framework of policies, processes, and standards that ensure the effective management of data assets
Data literacy	The ability to read, understand, create, and communicate data as information, including understanding data collection, analysis, interpretation, and presentation
Data migration	The process of moving data from one system or storage type to another
Data mining	The process of discovering patterns and relationships in large data sets
Data set	A collection of related data points or information used to train AI systems
Data steward	A person responsible for managing and overseeing an organization's data assets
Data visualization	The graphical representation of data and information using charts, graphs, and other visual elements
Database	An organized collection of structured information or data
Deepfake	Synthetic media where a person's likeness is replaced with someone else's using AI
Deployment	The process of making an AI system or application available for use
Deployment bias	When a system that works well in a test environment performs poorly when deployed in the real world
Ethical principles	Guidelines that ensure AI systems respect privacy, transparency, accountability, fairness, and human autonomy while promoting societal well-being
Explainability	The ability for humans to understand and trust decisions, recommendations, or predictions made by an AI systems
Feedback loop bias	When the output of an AI system influences future inputs, potentially reinforcing existing biases
Generative AI	AI systems capable of creating new content (text, images, code, etc.) based on patterns learned from training data
HIC (Human-in-Command)	A comprehensive oversight approach that considers broader economic, social, legal, and ethical impacts of AI systems
HITL (Human-in-the-Loop)	An approach where a human mediates all decisions made by the AI system

HOTL (Human-on-the-Loop)	An approach where humans monitor AI system operations and can intervene when necessary
Infrastructure	The hardware, software, networks, and facilities that support an organization's IT operations
Intellectual property rights	Rights that protect the investment of rights-holders in original content, including copyrights and accessory rights
KPI (Key Performance Indicator)	Measurable values that demonstrate how effectively an organization is achieving key objectives
LLM (Large Language Model)	AI models trained on vast amounts of text data that can understand and generate human-like text
Linguistic bias	When an AI algorithm favours certain linguistic styles, vocabularies, or cultural references over others
Machine learning	A subset of AI that enables systems to learn and improve from experience without explicit programming
Natural Language Processing (NLP)	The ability of computers to understand, interpret, and generate human language
Neural network	A computer system modelled on the human brain, designed to recognize patterns
Open source	Software whose source code is freely available for anyone to inspect, modify, and enhance
Participation bias	A form of sampling bias that occurs when certain groups choose not to participate in data collection
Pilot project	A small-scale preliminary study to evaluate feasibility, cost, and potential issues
Privacy	The principle that AI systems should respect and protect personal data and information
Prompt engineering	The practice of designing and optimizing inputs to AI systems to generate desired outputs
Proxy bias	When variables used as proxies for protected attributes introduce bias into the model
RPA (Robotic Process Automation)	Technology that automates repetitive tasks through software robots
Robustness	The ability of AI systems to maintain reliable operation and withstand adverse conditions or attacks
Sampling bias	When data is not randomly selected, resulting in a sample that is not representative of the population
Shadow AI	The unsupervised or unsanctioned use of AI tools within an organization outside of its IT and cybersecurity framework
Shadow IT	The use of IT systems, devices, software, or services without explicit organizational approval
Stakeholder	Any person, group, or organization that has an interest in or is affected by an AI project

SVG (Scalable Vector Graphics)	A web-friendly vector image format that can scale without losing quality
Temporal bias	When training data becomes outdated and no longer represents current realities
Traceability	The ability to follow and monitor the entire lifecycle of an AI system
Training data	The initial data set used to teach an AI system to perform its intended function
Transparency	The communication of appropriate information about AI systems in an understandable and accessible format
Use case	A specific situation or scenario where an AI system or application could be used
XAI (eXplainable AI)	AI systems designed to be interpretable and understandable by humans

INTER PARES
Parliaments in Partnership
EU Global Project to Strengthen the Capacity of Parliaments

This publication has been produced with the financial support of the European Union (EU), in partnership with the International Institute for Democracy and Electoral Assistance (International IDEA), as part of INTER PARES–Parliaments in Partnership, the EU's Global Project to Strengthen the Capacity of Parliaments.