WORLD
ECONOMIC
FORUM

Unlocking Public Sector AI

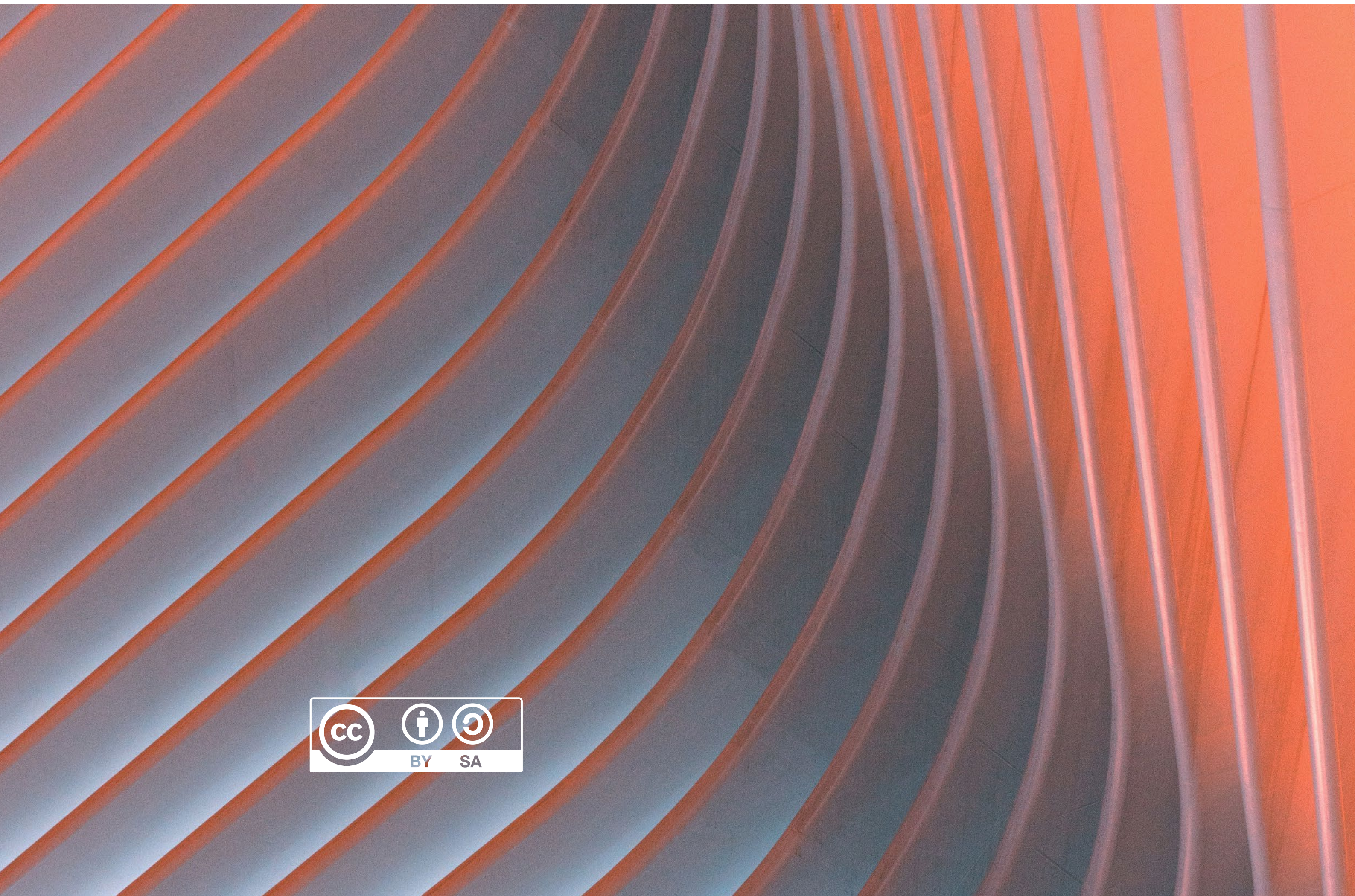
AI Procurement in a Box: Workbook

TOOLKIT

JUNE 2020

Cover photo credit:
Gus Moretta

Contents



A

AI risk assessment tool

A.1 Overview

This document sets out example decision criteria for conducting an artificial intelligence (AI) project risk assessment. An assessment of the potential risks involved in any solution that contains AI elements should be conducted as part of the planning phase of an AI procurement. This can also be a useful basis to develop a proportionate approach to AI procurement. It is important to approach AI procurement proportionally because not all guidelines as well as issues explored in guidelines apply to all procurement decisions in the public sector.

Purpose of this tool

The following table outlines some of the key questions you should consider when deciding your procurement strategy, choosing what requirements to include in your request for proposal (RFP) and assessing a solution. These questions have also been mapped to the issues that were set out in the *guidelines for AI procurement* document under the risk assessment header in the how to use the guide section.

How to use this tool

All these questions are designed to be answered with a yes or a no. Note that the list is not exhaustive and you should consider additional risks that are specific to your organization. For some of the questions below it might also be useful to consult the risk-based approach to AI adoption from the Canadian public sector, which divides AI systems developed at different levels. These categorizations provide insights into how to best approach AI procurement from a proportionality view and will help govern some of the decision-making.

Q1 **Is the solution intended for use in an area of public interest?**

If the project is within an area of intense public scrutiny (e.g. because of privacy concerns), interest and/or frequent litigation, then additional controls may be required. Fields such as health, social assistance, access and mobility, or decisions about permits and licences are examples of areas of applications that demand further consideration.

The higher the impact on individuals, businesses and communities, the more important it becomes to thoroughly consider AI ethics. The risk also increases when decisions made by the systems are linked to groups of people that are particularly vulnerable.

Q2 **Does the data used or generated by the solution contain any biographical or sensitive information?**

The more sensitive the data used or generated within an AI system the greater the number of checks you should build in.

Q3 **Are you comfortable with the data being stored and processed in an externally hosted solution?**

Consider whether the data has any protective markings or handling requirements that necessitate storage on authority infrastructure, such as a fully managed data centre or within a private cloud environment.

If your organization has a cloud-first policy and the data is suitable, a SaaS solution may be appropriate.

Q4 **Do you need to understand the details of how the data is being processed?**

For low-risk applications it might be appropriate to consider solutions that provide limited insight into how the data is processed, but if the solution is intended for processing personal information (such as medical applications), it may be useful to know the details of how it's been processed to ensure the outcome can be explained.

Q5 **Do you need the results of the processing to be validated by a human or is an automated output acceptable?**

If the output of the solution is intended for making critical decisions about services that are provided directly to citizens, then validation of the output is necessary. Alternatively, if you are considering a solution for managing cloud infrastructure to ensure the performance of a given application it might be appropriate for this to be fully automated.

Q6 **Do you have the skills and knowledge to define and assess the performance of the solution?**

Depending on the levels of expertise within your organization you may need to rely more heavily on a supplier or vendor to curate the solution for you. In this case you should expect the supplier to provide more detailed information about how they manage the solution.

If you have strong organizational data science skills, however, you should be able to more easily set the performance parameters, which makes custom solutions more achievable.

Q7 **Are you confident that the data intended for use in the solution is of good quality?**

The less sure you are about the quality of your data, the better it is to build in additional assurances to avoid bias.

Q8 **Are you happy for the supplier or vendor to enrich the data with external information as part of the processing?**

Some solutions will use external data feeds to draw conclusions from your data, and the source and utility of this external data should be considered when assessing what is acceptable for your organization.

The following table links the issues set out in the *guidelines for AI procurement* document to the most relevant questions.

FIGURE 1

Mapping guideline topics to the risk assessment tool

Issue	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Data		✓					✓	✓
Field of use	✓		✓	✓	✓			
Socio-economic impact	✓	✓	✓	✓	✓			
Financial consequences for agency and individuals	✓	✓	✓	✓				
Business function of the AI system			✓	✓	✓	✓		

A.2 Essential requirements in a proportionate approach

The following table outlines how the answers to the questions relate to the requirements described in the workbook Part C AI Procurement Specification and Evaluation Tool. It highlights the most important requirements related to the risk assessment. Please note that this does not mean that other requirements aren't also essential.

FIGURE 2 **How risk assessment relates to AI-specific RFP requirements**

		Essential requirements	Additional requirements
Q1	**Is the solution intended for use in an area of public interest?**	1.1	If **Yes**: Add more weight to 1.1
Q2	**Does the data used or generated by the solution contain any biographical or sensitive information?**	4.2, 4.3	If **Yes**: 4.4
Q3	**Are you comfortable with the data being stored and processed in an externally hosted solution?**	2.2	If **Yes**: 3.1, 3.2
Q4	**Do you need to understand the details of how the data is being processed?**	1.4, 1.7, 4.1	If **Yes**: 1.2, 1.3, 1.5, 7.1, 8.1
Q5	**Do you need the results of the processing to be validated by a human or is an automated output acceptable?**	1.6, 2.3, 9.1, 9.2	
Q6	**Do you have the skills and knowledge to define and assess the performance of the solution?**	3.3, 6.1, 9.3	If **No**: 3.4, 5.2, 5.3, 4.5, 10.1
Q7	**Are you confident that the data intended for use in the solution is of good quality?**		If **No**: 4.1
Q8	**Are you happy for the supplier or vendor to enrich the data with external information as part of the processing?**		If **Yes**: 2.1

A.3 Risk matrix

The risk matrix is designed to help the user determine their hosting and processing risks and what this means in terms of what types of solutions can be considered.

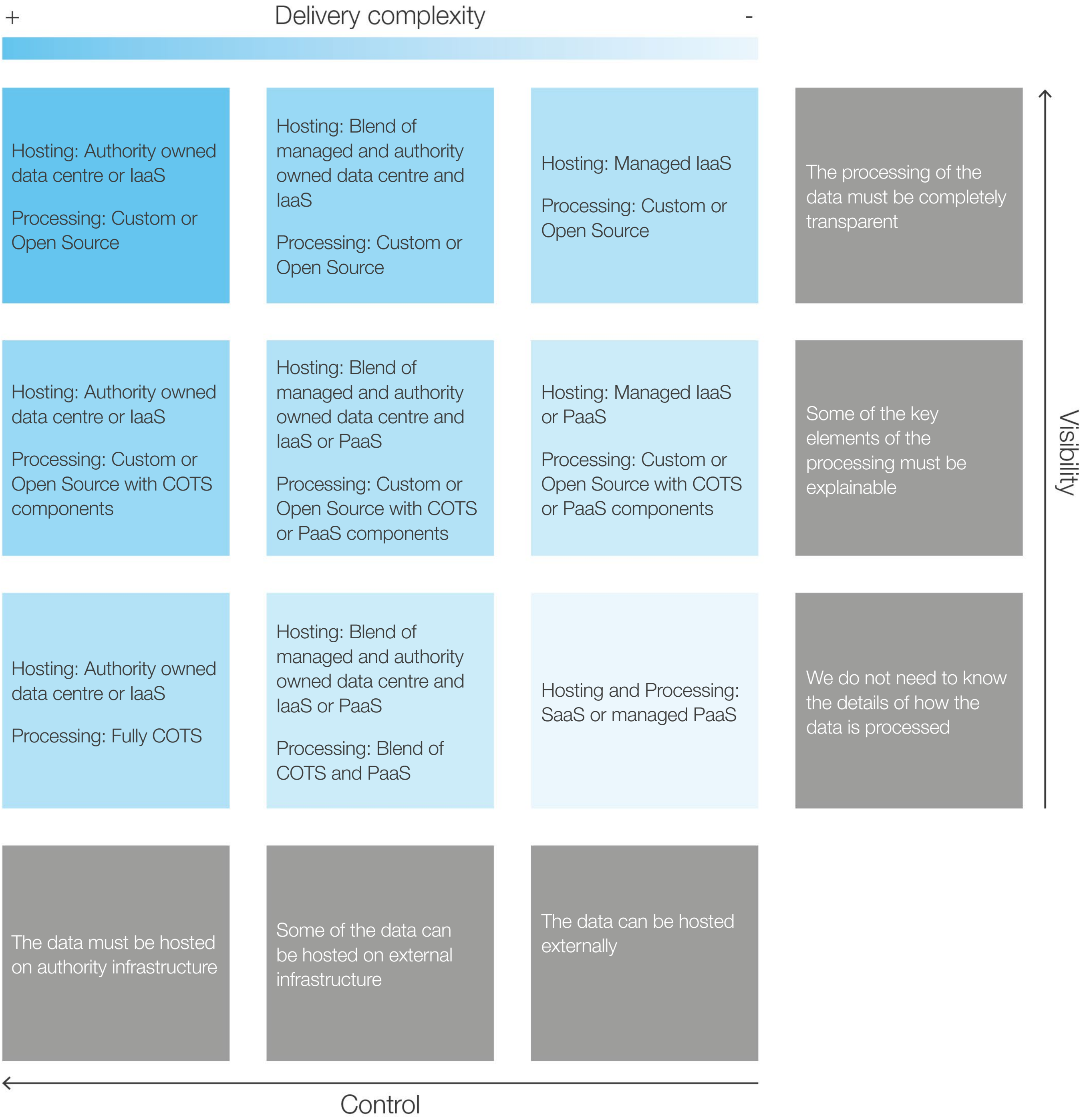
When considering the risks, you should:

- For control (or hosting) risks: consider your answers to questions 2 and 3 above.
- For visibility (or processing) risks: consider your answers to questions 4, 5 and 6 above.

Depending on your control and visibility posture the diagram will help you determine what solutions may be appropriate. For example, if all of the data can be hosted externally and you do not need visibility of the processing a SaaS offering could be appropriate. Note that for any box you land on from a visibility and control perspective, solutions that fit types above and/or to the left would also be appropriate, but bring a higher delivery risk.

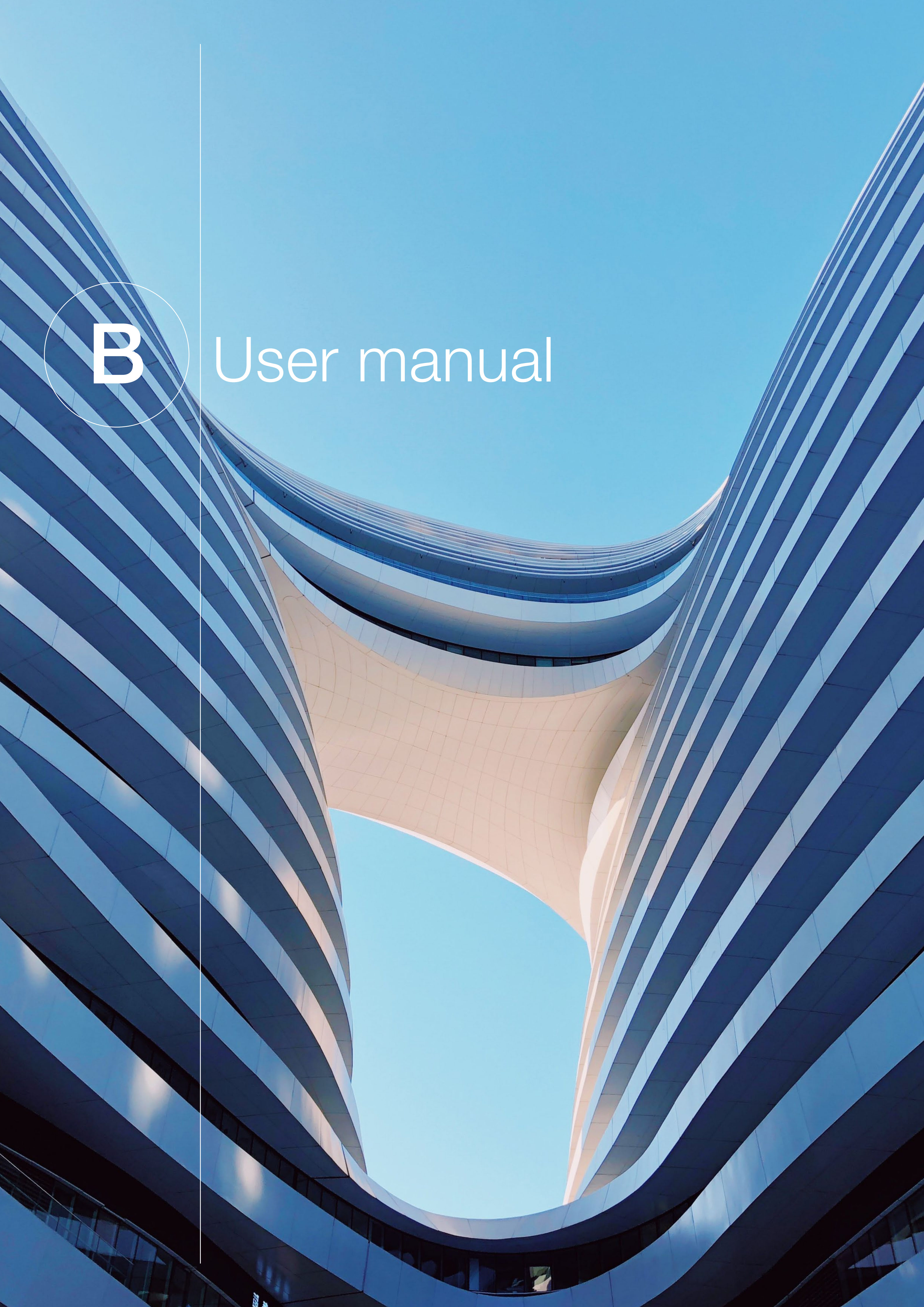
For clarity you can find definitions for Open Source[1], COTS[2], IaaS[3], PaaS[4] and SaaS[5] from the links found in the endnotes section.

FIGURE 3 **Risk matrix**



Hosting: Authority owned data centre or IaaS Processing: Custom or Open Source	Hosting: Blend of managed and authority owned data centre and IaaS Processing: Custom or Open Source	Hosting: Managed IaaS Processing: Custom or Open Source	The processing of the data must be completely transparent
Hosting: Authority owned data centre or IaaS Processing: Custom or Open Source with COTS components	Hosting: Blend of managed and authority owned data centre and IaaS or PaaS Processing: Custom or Open Source with COTS or PaaS components	Hosting: Managed IaaS or PaaS Processing: Custom or Open Source with COTS or PaaS components	Some of the key elements of the processing must be explainable
Hosting: Authority owned data centre or IaaS Processing: Fully COTS	Hosting: Blend of managed and authority owned data centre and IaaS or PaaS Processing: Blend of COTS and PaaS	Hosting and Processing: SaaS or managed PaaS	We do not need to know the details of how the data is processed
The data must be hosted on authority infrastructure	Some of the data can be hosted on external infrastructure	The data can be hosted externally	

B User manual



Overview:

Key factors to consider when beginning the procurement process for an AI-enabled solution

This manual provides a set of questions that highlight the main considerations that users should be able to address when implementing the guidelines.

1 Use procurement processes that focus not on prescribing a specific solution but rather on outlining problems and opportunities, and allow room for iteration.

Purpose of this tool

The user manual should help users to work through the different guidelines and find out how they apply to the specific project that they are working on.

How to use this tool

You can use the questions as a checklist at start of your procurement process.

1a. Make use of innovative procurement processes to acquire AI systems.

- Does your agency have access to a procurement vehicle(s) developed specifically for innovative technologies, such as AI?
- Have you engaged peers who have leveraged this procurement vehicle(s) in the past, whether inside or outside your agency, to learn from their experience?
- Are you leveraging any special mechanisms made available by the procurement process, such as agile procurement, challenge-based procurement, and/or dynamic purchasing systems?
- Does the procurement vehicle allow the procurement team to evalute responses within a reasonable amount of time, so as not to exclude potential participants?

1b. Focus on developing a clear problem statement, rather than detailing the specifics of a solution.

- Do you have a clear, concise problem statement that focuses on the needs of a user (e.g. benefit applicants)?
- Have you phrased your problem in a way that is technology agnostic?
- Have you engaged a group of peers and market partners, preferably knowledgeable in human-centric design, to confirm that you are addressing the root cause of the problem, as opposed to a symptom?

1c. Support an iterative approach to product development.

- Can you set expectations with providers through the RFP that the project must be delivered using an iterative (e.g. agile) approach?
- Can the problem be broken down into more manageable contracts and projects?

2

Define the public benefit of using AI, while assessing risks.

2a. Set out clearly in your RFP why you consider AI to be relevant to the problem and be open to alternative technical solutions.

- Do you have strong indications that AI is applicable to the problem? (e.g. do you have large amounts of data you could use to derive insights that address the problem?)

- Can the problem be addressed through a technology/solution that is likely to be better understood by the resources who will be responsible for delivering and operating it?

- Have you engaged peers and vendors to confirm that AI is a good solution to the problem?

2b. Explain in your RFP that public benefit is a main driver of your decision-making process when assessing proposals.

- Have you identified the protected groups, whether internal or external, who would be affected by the decision-making of the AI solution?

- Have you identified the potential biases that could exist in the data, which could unfairly affect the protected groups previously identified?

- Have you engaged the parties who will be affected by the tool and obtained their inputs (e.g. by holding citizen panels)?

- Have you identified success and failure criteria for the solution from the perspective of the stakeholders who would be affected by the solution?

2c. Conduct an initial AI risk and impact assessment before starting the procurement process, ensure that your interim findings inform the RFP and revisit the assessment at decision points.

- Have you identified the high-level potential impacts, including unanticipated consequences, that a solution could have on stakeholders? For example, for an AI-driven unemployment solution, could eligible recipients be wrongfully denied the benefit?

- Have you documented these potential impacts, together with viable mitigation strategies?

- Has executive management signed off the impact assessment?

- Have you included the results of the impact assessment in the RFP and asked vendors to suggest mitigation strategies?

3 Align your procurement with relevant existing government strategies and contribute to their further improvement.

3a. Consult relevant government initiatives, such as AI national strategies, innovation and/or industrial strategies and guidance documents informing public policy about emerging technologies.

– Have you identified relevant national strategies (e.g. AI strategy, digital strategy) and evaluated how your project can align?

– Have you identified and consulted on relevant policies and guidance frameworks, whether internal or external (e.g. innovation policies, technology policies, data policies and industry norms)?

3b. Collaborate with other relevant government bodies and institutions to share insights and learn from each other.

– Have you consulted peers, inside and outside your agency, who are specifically knowledgeable on govtech as well as the government's innovation and data policy agenda?

– Is there a public-sector community of practice or established body of knowledge that can be consulted for ideas on the solution and its potential benefits and risks?

– Have you consulted a repository of previous government AI projects for lessons learned?

4 Incorporate potentially relevant legislation and codes of practice in your RFP.

4a. Conduct a review of relevant legislation, rights, administrative rules and other relevant norms that govern the types of data and kinds of applications in scope for the project and reference them in the RFP.

– Have you consulted legal experts to ensure that the RFP addresses any and all legislation that could be relevant (e.g. with regard to privacy, national security)?

– Have you investigated whether there are commonly accepted industry practices regarding data?

– If applicable, have you established the governing law of data in cases of cross-border data flows?

– Have you set expectations in the RFP that contestability (i.e. the ability for a user to appeal against a decision made by the AI tool) will be built into the tool?

4b. Take into consideration the appropriate confidentiality, trade-secret protection and data-privacy best practices that may be relevant to the deployment of the AI systems.

- Have you agreed on what is commercially valuable information with the vendor to ensure that confidentiality and intellectual property protection are preserved?

- Have you consulted the freedom of information policies that would govern the required disclosures of information to the public to ensure accountability?

- Will the transfer and processing of personally identifiable data in relation to the solution be consistent with data protection and domestic privacy laws?

5 Articulate the technical and administrative feasibility of accessing relevant data.

5a. Ensure that you have proper data governance mechanisms in place from the start of the procurement process.

- How sensitive is the data that could be in scope? For example, could a solution potentially involve personally identifiable information (e.g., licence number, social insurance/security number, financial data, health data, etc.)?

- Are there processes in place to protect and manage data that could be used during the project?

- Are there processes in place to protect and manage data that could be used during the procurement process itself?

- Who will ultimately be accountable for the usage of data during the procurement process and the subsequent project (e.g. the Chief Data Officer, the data set's steward, etc.)?

- Is there an escalation mechanism for any procurement team members who may have a concern about potential data usage?

5b. Assess whether relevant data will be available for the project.

- Have you conducted a high level assessment to understand what data would be required to address the problem statement (e.g. necessary data sources or missing data)?

- Is the process to access this data understood, including identifying the data owner?

- Is there an understanding of how data would be accessed by the successful vendor(s) (e.g., onsite without leaving your data environment, remotely through VPN)?

5c. Define if and how you will share data with the vendor(s) for the procurement initiative and the subsequent project.

- Is there a case for sharing data with vendors (e.g. the benefits of sharing outweigh the risks)?

- If you have decided to share data, what mechanisms will you put in place to ensure the safety, confidentiality and privacy of the data?
- If you have decided to share data and you will be sharing a sample, how will you ensure the sample is representative of the users that will be affected by a possible solution?

5d. Ensure that you have the required access to data used and produced by the AI system.

- Have you asked for access to raw input, processed/combined and enriched data produced by the supplier(s) AI models?
- In case data sharing was not permitted, has the supplier been able to clearly articulate the reason for restricted sharing?
- Have you set out data ownership criteria for the AI system?

6

Highlight the technical and ethical limitations of intended uses of data to avoid issues such as historical data bias.

6a. Consider the susceptibility of data that could be in scope and whether usage of the data is fair.

- Would a solution use personally identifiable data, including but not limited to personal contact information, unique personal identifiers (e.g. licence number, social insurance/security number), financial data and/or health data?
- Would a solution use sensitive government data (e.g. military data)?
- What would be the impact of a data breach that could be in scope for the AI system?
- Does the data that could potentially be used for the project meet criteria for fairness, as specified in the guidelines?

6b. Highlight known limitations (e.g. quality) of the data in the RFP and require those tendering to describe their strategies on how to address the shortcomings. Have a plan for addressing relevant limitations that you may have missed.

- Does the team that owns and/or manages the data understand the data generation process?
- Have you consulted the data owner to obtain a high-level assessment of the integrity of the data?
- If data is of poor quality, have you considered alternative data sources, or consulted peers and/or market partners to seek advice on whether the data is usable and how much effort would be required to close the gaps?
- Is the data representative of the population to which the solution would apply or is the data biased? If biased, how will the bias(es) be addressed?

7 Work with a diverse, multidisciplinary team.

7a. Develop ideas and make decisions throughout the procurement process in a multidisciplinary team.

- Do you have a clear understanding of the skills that will be required to conduct the procurement process, including those relevant to policy, procurement, data and AI?

- Have you put together a team that has the skill set needed to effectively acquire and maintain the AI solution?

- How do research and consultations develop an understanding of the impact on diverse stakeholders/stakeholder groups?

- Is your team diverse? Does it promote inclusion in its composition? At a minimum, do you meet domestic laws of anti-discrimination?

7b. Require the successful bidder(s) to assemble a team with the right skill set.

- Will you require the successful bidder to include in its team resources with understanding of the affected group(s)?

- Will you require the successful bidder to meaningfully engage with the affected group throughout the design process of the solution?

- Does the RFP evaluation criteria assign a score for team diversity?

8 Throughout the procurement process focus on mechanisms of algorithmic accountability and transparency norms.

8a. Promote a culture of accountability throughout AI-powered solutions.

- Would the solution involve a human in the loop or would it be fully automated?

- Is the solution clearly understood by all stakeholders relevant to the RFP who would ultimately be accountable for the solution and its respective outcomes?

- Has an initial impact assessment for a possible solution been created as part of the procurement process, as well as been approved by the relevant stakeholders?

8b. Ensure that AI decision-making is as transparent as possible.

- Has an assessment been performed to gauge the necessary level of human oversight, given the sensitivity of the use case, the population affected by the solution and the data?

- Does the RFP ask the successful bidder(s) to create detailed user journey maps, including defining the level of information about the decision-making that the user would expect throughout the journey?

- Does the RFP ask the successful bidder(s) to provide users with an appeal mechanism when the user does not agree with an AI-driven outcome/determination?

- Does the RFP ask the successful bidder(s) to always inform users that they are interacting with a virtual agent, as opposed to a person?

8c. Explore mechanisms to enable interpretability of the algorithms internally and externally as a means of establishing accountability and contestability.

- Does the RFP require successful bidder(s) to provide documentation on the logic behind the algorithm, written in a way that can be understood by users with a limited knowledge of AI systems?

- Does the RFP require successful bidder(s) to provide detailed documentation of the solution and its processes?

- Does the RFP encourage successful bidder(s) to choose the least technically complex solution that will meet the requirements?

9 Implement a process for the continued engagement of the AI provider with the acquiring entity for knowledge transfer and long-term risk assessment.

9a. Consider during the procurement process that acquiring a tool that includes AI is not a one-time decision; testing the application over its lifespan is crucial.

- Has it been established whether the solution will be supported in-house or through a vendor? If through a vendor, will it be through the original vendor or a third party?

- Does the RFP require the successful bidder(s) to define how often the model should be updated to maintain the required performance?

- Does the RFP require the successful bidder(s) to agree to third-party solution audits and to provide the necessary level of access required for maintenance and support?

- Does the RFP ensure the necessary level of access, interoperability and data portability required for maintenance and support?

- Have you defined whether the optimal way to source the solution is through one or multiple contracts (e.g. through consideration of budget, risk management, access to skills)?

9b. Ask the AI provider to ensure that knowledge transfer and training are part of the engagement.

- Does the RFP require the successful bidder(s) to define how often and by whom the model should be updated to maintain the required performance?

- Does the RFP require the successful bidder(s) to define how they will team up with the public-sector authority to share insights into the technology and provide knowledge transfer?

- Does the RFP require the successful bidder(s) to provide thorough and holistic documentation about the solution?

9c. Ask the AI provider for insights into how to manage the appropriate use of the application by non-specialists.

– Does the RFP require the successful bidder(s) to provide training material and/or documentation sufficient for relevant non-technical staff to be able to effectively operate and govern the solution?

– Have you incorporated access control mechanisms to prevent unauthorized and unintended uses of the solution?

9d. Make ethical considerations part of your evaluation criteria for proposals.

– Does the RFP ask bidders to provide their own ethics framework for data and AI?

– Does the RFP require bidders to comply with existing government ethics standards, including those created specifically for AI?

– Does the RFP ask bidders to propose process and/or system metrics that reflect a consideration for ethical standards?

– Does the RFP's scoring assign non-trivial weight to ethics capabilities and experience shown by bidders?

10 Create the conditions for a level and fair playing field among AI solution providers.

10a. Contact a variety of AI solution providers in various ways.

– How could traditional and non-traditional partners, such as start-ups and academia, add value to the project?

– Have you actively sought new ways of market engagement, such as hosting a Q&A session, pre-RFP sessions to discuss the problem, supplier days, hackathons or co-working space presentations?

10b. Engage vendors early and frequently throughout the process.

– Have you validated the problem statement and your assumptions (e.g. user needs, applicability of AI) with potential partners?

– Have you defined a single point of contact for bidders who have questions and provided the relevant contact information?

10c. Ensure interoperability of AI solutions and require open licencing terms to avoid vendor lock-in.

– Does the RFP set expectations that tools used be open source and that open standards be leveraged as much as possible?

– Is there a clear understanding between vendors and the contracting agency regarding IP ownership of the project's deliverables?

– Does the solution involve technologies that contain patents or other intellectal property and if so is licencing available royalty-free?

C AI specification and evaluation tool

Overview

This tool provides examples of requirements for civil servants to include in a request for proposal (RFP). It also highlights examples of robust AI systems development as well as deployment practices to look out for in the responses or discussions with suppliers. It is intended to be used during the procurement process in conjunction with the AI procurement guidelines as well as the risk assessment that should allow for a proportionate approach to procurement. The key principle for AI procurement is to clearly describe the problem the contracting authority is aiming to address, focus on outcome-based criteria and not overspecify the AI system, ensuring that the most suitable system is purchased and to innovation is supported.

Purpose of this tool

This document aims to provide you with an introduction on what to consider when evaluating AI systems during the procurement process. It gives examples of several questions that you can ask while procuring AI systems from suppliers in categories such as intended use, accuracy of data, fairness and transparency of algorithmic-based decision flows, data security and effectiveness of the systems in meeting intended use.

How to use this tool

You can consult this document while drafting RFPs and evaluating responses. To use this document effectively please refer to the AI risk assessment tool in the workbook to identify which AI systems and procurement considerations may be more relevant for your project and to assess your requirements.

This document does not aim to provide a recommendation for an exhaustive list of necessary requirements that suppliers need to respond to. It highlights issues that can be considered when setting out specifications in RFPs or evaluation responses in an iterative process. You might already have robust processes in place for some of the issues mentioned below. These examples should not replace those processes, but rather introduce additional criteria to consider due to the complexity added by the AI system. The table below outlines how to use the document in more detail.

Note that the requirements and criteria in this document are for guidance purposes only. It is essential that you consider the importance of the requirements against your needs and tailor your questions and evaluation accordingly.

1 **Purpose:** The supplier understands the problem to be solved and the purpose and goals of the technical AI system

Sample specification	Key considerations to look out for in the answers
1.1 Describe the area of the problem space that is addressed by your AI system.	1. Does the supplier articulate the part of your problem that is addressed by the AI system? 2. Does the supplier recognize and describe any limitations of the AI system for the problem at hand? 3. Is it made clear if the AI system is dependent on those AI elements being added? 4. Can the supplier justify why use of AI/ML is the best approach to address the problem?
1.2 Is your approach built on an existing AI system (Commercial Off the Shelf (COTS)) or will it be custom-made or a mix of the two?	5. Does the supplier describe the elements of the AI system and where they originate?
1.3 Describe what algorithms or techniques you anticipate the AI system to implement.	6. Does the supplier explain the techniques applied in the AI system, including use of any algorithms and associated software libraries for the algorithms? 7. Can the supplier explain how the system operates in an easy to understand way for various audiences?
1.4 Describe the approach to ensuring that use of AI is necessary and proportionate in the AI system.	8. Does the supplier explain the metrics and evaluation methods used and how they have impacted the selection of data that will be used in the proposed AI system? 9. Can the supplier articulate potential risks of using the AI/ML solution and risk mitigation strategies?
1.5 Describe how you have ensured that the AI system is proportional to the data available.	10. Does the supplier explain how it will be ensured that data needs required to produce the intended outcome are considered proportional? 11. Is the supplier capable of mitigating the data supply that they need from the operator? 12. Does the supplier explain the need to access various data sets?
1.6 Explain how all end users have been considered throughout the design and implementation process.	13. Does the supplier describe how the proposed AI system supports transparency and explainability characteristics not just for the data subject, but the end user/operator as well? 14. Does the supplier set out a plan that allows for user testing and an iterative design approach and risk mitigation?
1.7 Explain how you will demonstrate accountability for the goals and outcomes of the AI system.	1. Does the supplier describe the end user training they commit to deliver to ensure the ongoing health and maintenance of the AI system and outcomes? 2. Is the supplier providing documentation detailing how the AI system can be configured or adapted if the results are not delivering the goals or the AI is not acting in an ethical or understandable manner?

2 **Consent and control:** The developer will ensure that they have consent from the data subject before processing data or training an algorithm, and that human operators can control the outcome

Sample specification	Key considerations to look out for in the answers
2.1 Please provide evidence that you have considered the legal and ethical implications and gathered consent for processing and capturing the data throughout the full lifecycle of the AI system. NOTE: criteria correspond to COTS AI system. Same criteria can, however, apply to tailored products (e.g. “The supplier provides information on what individuals will be told, when they will be made aware, what kind of consent will be needed from them, and what the procedures will be for gathering consent.”).	17. Can the supplier articulate how it was decided whose data to use or about whom to make inferences? 18. Is it clear that data subjects know that their data is being used or that inferences are being made about them? 19. Does the supplier provide information on what individuals were told, when they were made aware, what kind of consent was needed from them, and what the procedures were for gathering consent? 20. Does the supplier highlight potential risks to these individuals or groups and how the service output might interfere with individual rights? 21. In the case of risk identification, does the supplier describe how the risks are being handled or minimized? 22. Does the supplier describe how the rights of individuals who provided the data were safeguarded throughout the process? 23. Is it made clear whether individuals have the option to withdraw their data and opt out from inferences being made about them? If yes, what is the withdrawal procedure? Suppliers should ensure that all raw input, processed, training and enriched data is accessible and usable in a timely manner for the public-sector authority, especially for monitoring and inspection. Ideally the suppliers process and data governance should make sure that persistent ownership and access to this data is granted to the public-sector authority, including third party and/or open source data sets.
2.2 Describe your approach for allowing access and control of the data within the AI system.	24. Does the supplier provide access to the AI model(s) input data, including any third party or open source data including mechanisms for controlling the flow of data? 25. Can the supplier provide access to all the AI-model(s) training data and when this is not feasible explain the process for providing a representative sample? 26. Can the supplier provide full access to the AI model(s) processed/combined and enriched data (i.e. key features, inferred scores/metrics) and when this is not feasible explain the process for providing a representative sample? 27. Does the supplier describe the level of contractual ownership that will be granted to the above data and for what period?
2.3 Describe the level of human decision-making at critical control points.	28. Does the supplier describe the approach to active monitoring to track user behaviour to identify irregular patterns that may indicate signs of unintended consequences? 29. Does the supplier mention operational bias reviews to track model inputs and outputs to identify irregularities that may indicate bias? 30. Does the supplier mention that they might retrain the model in agreement with the operator using new or more up-to-date data to account for changes in user behaviour?

3 **Privacy and cybersecurity:** The supplier will not introduce harm through unintended consequences or poor practice

Sample specification	Key considerations to look out for in the answers
3.1 Describe your privacy and cybersecurity approach for the proposed AI system as well as how the data will be protected. NOTE: COTS and bespoke AI systems will have dependency on security controls managed by the authority.	31. Does the supplier deploy well-established techniques, security processes and standards to protect the data, for example, encryption and anonymization, where appropriate and feasible? 32. Does the supplier describe how need-to-know principles for data access are applied and the decision criteria for allowing access to data and AI models? For legitimate and logical reasons, protected and or sensitive data may be required and processed by the AI system. Development teams should invest time in understanding the reasons why the data is sensitive and the impact on the data subjects in the event of a biased decision or data breach. Typically, AI systems must not be designed to be fully autonomous. Human operators or even data subjects should be able to intervene or interrupt in the event of incorrect or harmful decisions being made and/or be asked to confirm a processing phase or learning step before it commences.
3.2 Describe the potential threats to the system or AI system from external or internal adversaries. NOTE: Bespoke AI systems may have dependencies on authority risks, but should be able to describe risks that are specific to the AI system.	33. Does the supplier define how the system could be attacked or abused? 34. Suppliers could: – List applications or scenarios for which the service is unsuitable. – Describe specific concerns and sensitive use cases and what procedures can be put in place to ensure that the service will not be used for these applications, or if the service needs to be used in a sensitive use case the precautions being taken to mitigate harm. – Underline that they will verify AI model stability when exposed to sub-system compromise and/or outages. – Describe how they are securing user or usage data. – Identify if usage data from service operations is retained and stored. – Ascertain how the data is being stored and for how long the data is stored. – Mention how they will verify if enriched and/or inferred user or usage data is being shared outside the service and who has access to the data. – Describe how the service checked for robustness against adversarial attacks, including once it is integrated/deployed at scale. – Explain how robustness policies will be checked and the type of attacks considered. – Propose a plan to handle any potential security breaches based on accepted industry best practice.
3.3 Explain your test processes, including the specialist expertise used to assess the AI system.	35. Does the supplier provide evidence that the AI system has been tested and that AI domain experts were involved in the development, testing and deployment? 36. Can the supplier describe how the AI model(s) will be monitored and checked to highlight potential malicious manipulation (internal and external)?
3.4 Please provide evidence of previous case studies of where the AI system has been implemented and how the output has been interpreted, highlighting best practice.	37. Does the supplier provide evidence of where the AI system has been used before? 38. Can the supplier point to previous use cases that include description of how the output has been consumed, drawing out if any harm or negative impact on the end users or data subjects was introduced through misuse or misinterpretation?

4 **Ethical considerations:** Will the service or AI system be fair in its decision making and processing

Sample specification	Key considerations to look out for in the answers
4.1 What data limitations have you identified and what strategies will you implement to address these data limitations? NOTE: this is applicable only when the authority has shared data with the supplier or when the supplier is using pre-trained models or their own data. Otherwise, this should be assessed during AI system design.	39. Can the supplier describe where they have missing or poor quality data? Are they able to identify potential risks that arise from missing or poor data and can they articulate how they are mitigating these risks? Suppliers should be able to describe how data bias policies will be checked (with respect to known protected attributes), bias checking methods and results (e.g. disparate error rates throughout different groups). Suppliers should also be aware of the personal or unconscious bias inherent in the development team and the human operators of the AI system and how it influences the output of the system. Bias may also be a legitimate input in certain problem sets or use cases, but unconscious or personal bias that undermine the correctness of the outcome or introduces harm must be avoided. There needs to be a focus on detecting unconscious or personal bias during the training and testing of the algorithm. Given the needs to adapt processes to ensure fair treatment for persons with disabilities as employees and as service users and citizens accessing government information and services – suppliers must be required to demonstrate that the end-to-end process they are influencing or managing is non-discriminatory – it is I important, but far from sufficient, to just address data bias.
4.2 How will you ensure that the AI system fits the requirements of data ethics frameworks and policies prior to going live?	40. Is the supplier able to demonstrate how data ethics principles referred to in the RFP are considered in designing, building and supporting their AI system?
4.3 Describe the approach to eliminate (or minimize) bias, ethical issues or other safety risks as a result of using the service.	41. Can the supplier describe the possible sources of bias or unfairness assessed and where they arise from – the data, the techniques being implemented or other sources? 42. Is there any mechanism for redress if individuals are negatively affected?
4.4 Describe the process for ensuring that the development team adopts an ethical mindset.	43. Does the supplier offer training or have an awareness process to ensure their team understands the potential impact of creating an AI system that produces an incorrect, biased or disproportional output? 44. Can the supplier describe how they educate their staff to understand and accept that individuals have unconscious bias and understand their responsibility for ensuring this does not affect the operation of the AI system?
5.5 Explain how the AI system will be tested during the life cycle to detect bias and the remediation steps if it is introduced.	45. Can the supplier describe bias policies models and bias checking procedures, as well as how they will monitor and verify results (e.g. disparate error rates throughout different groups) with a focus on controls for unacceptable bias and/or defined thresholds? 46. Does the supplier highlight life cycle considerations and maintenance of the AI system? Do these considerations include model validation processes to assess performance against defined tolerances and/or thresholds and demonstrate their ability to highlight other potentially less visible problems (i.e. overfitting)?

5 **Explainability:** Can the supplier adequately explain how the AI system functions to the affected consumer, data subject or operator

Sample specification	Key considerations to look out for in the answers
5.1 Describe the provisions in the AI system to ensure that the outputs are explainable and/or interpretable.	47. Is the supplier able to define how their organization approaches ethics? 48. Is the supplier able to show how they aim to aid the explainability of their AI system (e.g. directly explainable algorithm, local explainability, explanations via examples)? 49. Can the supplier provide clear guidance and explanations on how the results of the AI process should be interpreted? 50. Does the supplier outline the target user of the explanations (AI expert, domain expert, general consumer etc.) and ask them to describe any human validation of the explainability of the algorithms? 51. Does the supplier highlight key parameters and inputs to their AI model(s) and how they affect the outputs (i.e. sensitivities)?
5.2 Would you allow independent, third party audit(s) of the AI system? If your answer is no, please explain.	52. Is the supplier able to allow for external audits? 53. In the case that an external audit is not possible, justification must be provided.
5.3 Describe how you enable end-to-end auditability of the AI system.	54. Can the supplier describe what information is captured throughout the AI system and provide a taxonomy to describe the meaning of the information? 55. Is the supplier able to provide documentation related to the development and support of the AI system, for example, test reports, logs and quality criteria?

6 **Concept drift:** Ensuring the system does not drift from its intended purpose

Sample specification	Key considerations to look out for in the answers
6.6 Explain how you will ensure the AI system or service does not drift from its intended purpose or outcome.	56. As algorithms are learning continuously after they are developed it is possible for them to drift from the original concept and deliver different results. Providers can be assessed on their approach to the following: – What is the expected performance on unseen data or data with different distributions? – Does the system make updates to its behaviour based on newly ingested data? – Is the new data uploaded by users? Is it generated by an automated process? Are the patterns in the data largely static or do they change over time? – Are there any performance guarantees/bounds? – Does the service have an automatic feedback/retraining loop or is there a human in the loop? – How is the service tested and monitored for model or performance drift over time? – Is the supplier providing performance drift monitoring KPIs that prompt retraining if there are any unexpected changes? – How can the service be checked for correct, expected output when new data is added? – Does the service allow for checking for differences between training and usage data? – Does it deploy mechanisms to alert the user of the difference? – Do you test the service periodically? – Does the testing include bias or fairness related aspects? – How has the value of the tested metrics evolved over time?

7 Interoperability and other standards

Sample specification	Key considerations to look out for in the answers
7.1 Explain how your system or service conforms to specific international or local open interoperability standards or other relevant standards relating to cyber security, coding quality, safety, testing, accessibility and usability. Examples are the IEEE standards as well as GDPR for personal identifiable information (PII).	57. Does the supplier explain how the AI elements of the system or service operate with the following? – Required data storage/access requirements? – Operational monitoring/compliance tools? – Standard system elements, including COTS, Operation support systems (OSS) and/or custom? 58. Can the supplier demonstrate the range, velocity and veracity of data and features that can/will be provided for wider potential use/developments? – Detail interfaces (i.e. API) and integration dependencies (particularly OSS or custom elements)? – Provide an approach for future interoperability requirements? 59. Does the supplier include business continuity management measures such as documentation and access to key processes and algorithmic steps for the AI model(s), where these are not provided as part of the normal delivery of the AI system?

8 Due diligence on existing algorithms or COTS AI systems

Sample specification	Key considerations to look out for in the answers
8.1 Describe the architecture of the AI system, including use of external COTS or open source elements and the function they provide in the AI system. This should consider the data used by each element of the AI system and how the output of that element was validated.	60. If an AI system is based on an existing algorithm or will integrate with another functionality, the supplier should be able to describe the full nature of the system. For example, a COTS AI system could introduce unknown ethical risks if used improperly. Potential areas for consideration could be: – Is the service or AI system based on COTS, OSS and/or legacy AI system(s)? – Which datasets was the service trained on? – Were there any quality assurance processes employed while the data was collected or before use? – Were the datasets used for training built for purpose or were they repurposed/adapted? – Were the datasets created specifically for the purpose of training the models offered by this service? – Are the training datasets publicly available? – For each dataset: Does the dataset have a datasheet or data statement? – Did the service require any transformation of the data in addition to those provided in the datasheet? – Was synthetic data used and how was this generated? – How were the models trained and when were they last evaluated for correctness? – How often are the models retrained or updated? – Did you use any prior knowledge or reweight the data in any way before training? – How is testing conducted by the service provider? – Which datasets was the service tested on (e.g. links to datasets that were used for testing, along with corresponding datasheets)? – Could these datasets be used for independent testing of the service? Did the data need to be changed or sampled before use? – Please provide details on train, test and holdout data and what performance metrics were used (e.g. accuracy, error rates, AUC, precision/recall)?.

9 Lifecycle management

Sample specification	Key considerations to look out for in the answers
9.1 Explain how you will ensure the AI system or service does not drift from its intended purpose or outcome.	61. Is the supplier able to provide information on any existing training courses or documentation they have available? 62. Does the supplier include the creation of training materials as part of their offering bespoke AI systems?
9.2 Explain how you will ensure usability for non-trained staff.	63. Can the supplier describe the target user for the AI system, including expectations around their skills? 64. Can the supplier articulate how users can be trained to use and understand the AI/ML solution being implemented? 65. Can the supplier outline the types of skills required to support or use the AI system and the role types they would expect to see? For example, system admin, data scientist, end user.
9.3 Explain how the AI system will be maintained, how its accuracy and integrity will be sustained over time, and whether third party providers could be engaged for these activities.	66. Is the supplier able to describe the handover process in the case of a bespoke or COTS offering? This should detail: – Accuracy metrics and thresholds to ensure the integrity of the AI system. – Maintenance processes and activities. – Support contracts. – Suitability for third party support. 67. Is the supplier able to provide a service agreement detailing the approach to AI in case the system is based on software as a service (SaaS)? 68. Can the supplier demonstrate scale deployment considerations for their AI model(s) (e.g. limit to data coverage, minimum model training requirements, system processing time sensitivities, etc.)?

10 Skills

Sample specification	Key considerations to look out for in the answers
10.1 Can you demonstrate how you will assess the competencies, qualifications and diversity of the team that will develop and deploy the AI system?	69. Can the supplier outline how they are drawing on appropriate skills to be domain experts in the field of AI and in the area the AI system is to be applied? 70. Do the supplier skills set match standards referenced in the Skills Framework for the Information Age (SFIA framework)? [6] 71. Does the supplier highlight the importance of diversity in AI development and explain how this is considered in the composition of the delivery team and provide strategies to increase diversity in AI development if diversity requirements cannot be met by the immediate team?

D How to kick-start the implementation of the guidelines

World Economic Forum AI procurement workshop templates

↓ Day One

↓ Day Two

↓ Day Three

E Case studies

Below: Guwahati, India

1 Case study India

Controller General of Patents, Designs and Trade Marks

“ The aim is to enhance efficiency, uniformity and consistency.

Objective

The Indian Controller General of Patents, Designs and Trade Marks (CGPDTM) wanted to make use of artificial intelligence (AI), blockchain, internet of things (IoT) and other new technologies for its patent processing system. The aim is to enhance efficiency, uniformity and consistency within issues ranging from inception of a possible IP to its enforcement.

Why AI?

The patent processing system is a manually extensive and long process. As such, AI was considered a potential solution to modernize, automatize and strengthen the transparency of the process. It is also hoped that having a stable and efficient IP regime in the country encourages innovation to achieve the country's industrial and economic development goals. The initiative was part of a larger government effort to explore the use of blockchain and AI in diverse areas such as education, healthcare, agriculture, electricity distribution and land records.

Background

The CGPDTM is responsible for administration of all major IPR legislations in the country regarding patents, designs, trademarks, geographical indications, copyrights and semiconductor integrated circuits layout-design. The office processes approximately 55,000 applications per year.

Action

The procurement process was divided into two phases – the initial expression of interest (EOI) and request for proposal (RFP). The EOI was made available publicly on an existing e-tendering platform well-known to the business sector. The aim was to seek proposals as to how best to shortlist vendors for the purpose of hosting a limited tender. The participation of small and medium-sized enterprises was greatly encouraged through lower eligibility standards.

The agency suggested different areas for the proposals (electronic data processing, screening, prior art searching, pre-grant opposition etc.) and companies were invited to pitch various solutions and technologies. The selection criteria for the EOI was based on track-record for similar projects, general qualifications of key staff, financial strength and accreditation and certifications. Hence, the agency ensured that the vendor had the right skills set to develop and deploy the AI solution by demanding proof of certifications, references and past experiences.

The RFP evaluation was much more focused on a specific type of solution and was based on technical bid evaluation, technical demonstration and financial bid. For the financial bid, the lowest bid was considered successful. Throughout the process, vendors were invited to submit queries for specific questions, which were answered at specific moments and made publicly available. It was agreed that the solution developed and furnished belongs exclusively to CGPDTM. The vendor had to grant a non-exclusive licence to access, replicate and use the application software, the custom software and any proposer owned software embedded in the systems.

Ethical considerations

An important consideration for the deployment of the solution was the explicability of the search queries and the avoidance of biases. This was ensured by making the source code of the solution available to the public. The RFP also made clear that any sensitive data provided would be hosted either on premises or through an API access[8] and would only be available to the successful vendor for testing/development phase. Furthermore, it was clarified that no data would be hosted outside India.

Lessons learned: Which guidelines were harder to implement?

"Support an iterative approach to product development."

"Assess whether relevant data will be available for the project."

The "Eligibility and Financial Criteria" methodology used to select a vendor was hard to understand for many RFP participants. One aspect that led to confusion was the required accuracy of 75% for developed models. The RFP did not give a clear definition of "accuracy" and did not provide historical data for training and testing of the models. As machine ML/AI models improve accuracy over time as they learn and get better, it was hard for the RFP participants to develop a 75% accuracy without access to relevant data. In addition, this evaluation criteria lacked transparency and didn't support an iterative approach to product development. Following the concerns raised by the participants, the CGPDTM lifted that requirement.

"Develop an understanding of the skills that are needed to effectively acquire and maintain an AI-powered solution, before starting the procurement process."

Successfully designing and deploying AI in an organization as big and complex as the CGPDTM was a major technical and human challenge. Assembling a team with experience in change management and technical expertise on integration with existing software and datasets could have helped to better navigate the procurement and implementation process.

Success factors: Which guidelines were successfully implemented?

"Aim to include your procurement within a strategy for AI adoption across government and learn from others."

This project was part of a larger government of India-wide effort to adopt and enhance the use of latest technologies and as such, senior government functionaries were very active in making the procurement process a success. This strong leadership from the government ensured that the right resources were employed and the process moved forward.

"Reach out in various ways to a wide variety of AI solution providers."

"Create the conditions for a level and fair playing field among AI solution providers."

While providing opportunities to various firms to compete, the public EOI also boosted innovation and the diversity of the proposed solutions. Newly established providers were also given the opportunity to compete for this public-sector contract through lower requirement standards.

"Focus on developing a clear problem statement, rather than on detailing specifications of a solution."

An extensive and clear description of the IPO workflow and use-cases for AI made it easy for participants to identify opportunities. Documenting user needs and challenges for each stage of patent applications was crucial for AI system providers to understand the problem.

"Define if and how you will share data with the vendor(s) for the procurement initiative and the subsequent project."

The RFP was clear on data governance during and after the procurement initiative. The governance approach specified who would be granted data access, the purposes for which a vendor would be authorized to use the data and the minimum requirements for hosting/reviewing the data.

"Require the successful bidder(s) to assemble a team with the right skills set."

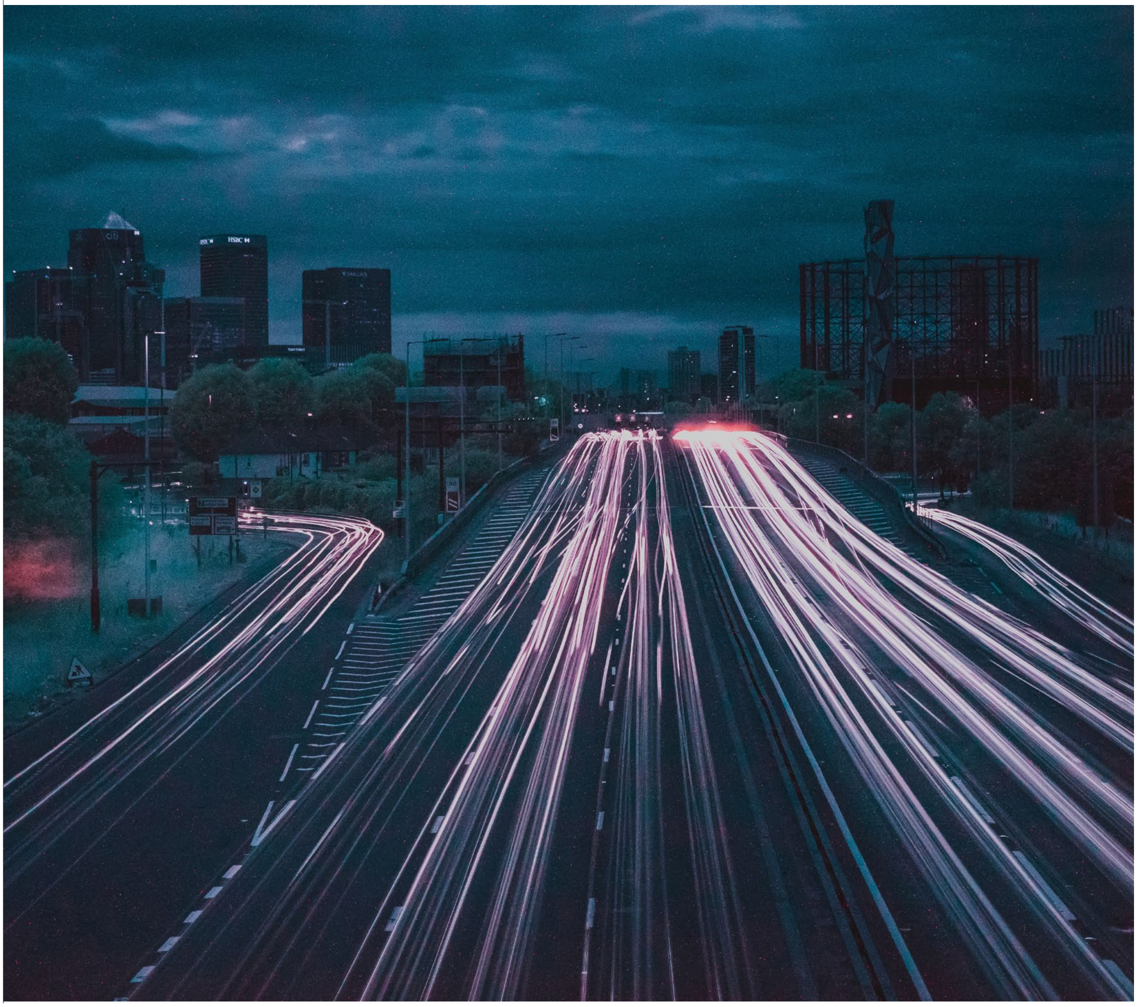
Evidences of skills and qualifications of key team members were required in the initial EOI. Evidence of bidder's resources for deploying the solution were also assessed and were part of the decision-making criteria.

Below: Bugsby's Way,
London, United Kingdom

2

Case study United Kingdom

Driver & Vehicle Standards Agency



Objective

The Driver and Vehicle Standards Agency (DVSA) wanted to make use of digital technologies to ensure that vehicle standards are enforced while at the same time saving time and costs. A data-driven approach should help the agency to conduct intelligent inspections of authorized garages conducting the vehicle standards test.

Why AI?

The team held a lot of data that it couldn't use effectively. The testing was resource intensive and the previous process did not allow for targeted inspections. Clustering techniques offered insights that were previously not available. This helped to make predictions that now support a more targeted approach to inspections.

Background

> “ The department only became aware of the power and opportunities of applying AI when it received the responses to the invitation to tender – and, at a more detailed level – once it started working with the partners.

The DVSA is an executive agency of the United Kingdom Department for Transport, which among other things supervises the MOT scheme, a vehicle standards examination, ensuring that authorized garages carry out tests to the correct standards. This examination, referred to as “the MOT”, assesses vehicle safety, roadworthiness and exhaust emissions and is required in the UK for most vehicles over three years old and used on anything that can be classified as a road. Each year, 66,000 testers conduct 40 million MOT tests in 23,000 garages. The inspection of the authorized garages was resource intensive and the knowledge was limited to effectively target inspections of these garages.

The DVSA made the decision to further invest in the MOT to improve the service in a number of ways, including quality of the service to the end user (motorist), test quality, reduce fraud risks and improve efficiency. The DVSA had insufficient capacity to do this so chose to procure two digital partners. As well as delivering some of the improvements (in consort with DVSA as part of blended agile teams) the partners would also develop the department's in-house skills.

The DVSA released an invitation to tender (ITT). The AI aspects of the work were part of this larger contract for digital transformation and the department only became aware of the power and opportunities of applying AI when it received the responses to the invitation to tender – and, at a more detailed level – once it started working with the partners (as part of options for solving business challenges).

Action

During the procurement process the DVSA ensured that the ITT set out clearly what challenges it wanted to solve and what outcomes it sought. The DVSA used the Digital Outcomes and Specialist Framework, which is a framework agreement that focusses on the digital transformation of public sector services.[9] The ITT did not ask for AI as a technology, but laid focus on the use of technologies that would deliver the most effective outcome. The aim of the procurement effort was to contract digital services and skills that would help the team to identify and deploy the right tools and systems to address the delivery challenges, in particular improving the DVSA inspection of authorized garages that conduct MOT tests. During the ITT stage, pricing arrangements were kept simple with partner effort paid on a time and materials basis at agreed rates. It was required that all IP would be owned by the DVSA.

The project started with a set of mini discoveries, which enabled the agile nature of the work. These covered a number of areas and included the following:

- Improving MOT test quality through better supporting testers
- Better enabling the DVSA to know which garages presented the greatest risks of testing poorly
- Identifying those applying to be involved in MOT that may present risks to the integrity of the MOT service

In collaboration with the supplier, the DVSA applied a clustering model against garage test data from a three-month period.[10] The clustering model grouped MOT-authorized garages based on the behaviour they show when conducting MOT tests, such as the test duration, time of test and result of inspection (against expected). The DVSA created a risk (of testing incorrectly) score for each garage, which allowed the department to rank garages and their testers and helped it identify regional trends. The model was validated against those who had been identified as doing things incorrectly, ensuring that the model could learn what behaviours were good indicators of wrong-doing.

An important consideration was the ability to explain the model and the human in the loop. It is important to explain the outcome of the risk rating without losing the integrity of the test. Having a human in the loop who interrogates and decides to take action on the risk score was crucial to make the use of AI successful. All the data used for the AI system was data that was already collected by the DVSA and it did not include a great amount of sensitive data. Suppliers had visibility of some data, but not off-site access.

The lifecycle management of the tool was not fully factored in upfront and became a challenge once the technology was developed. The DVSA team identified this as an issue and worked with suppliers to put together a plan to bolster the skills of the department's continuous improvement team. This ensures that the system continues to work effectively and meets users' needs, as well as technical support that addresses issues related to hosting and live service failures.

Impact

50%
the fall in examiners' preparation time for enorcement visits

The DVSA can now target its resources at the garages and testers with the highest risk score. By identifying areas of concern in advance, the examiners' preparation time for enforcement visits has fallen by 50%.

There has also been an increase in disciplinary action against garages, meaning standards are now being better enforced. As more garages are delivering better MOT standards, there are more cars on the road that comply with roadworthiness and environmental requirements.

Lessons learned: Which guidelines were harder to implement?

“Support an iterative approach to product development.”

It was important to find the right balance between agile delivery and the focus on price in the evaluation of the proposals. Since prices and timelines might shift due to the agile nature of the work, you must ensure that you reflect this in the scoring of the invitation to tender and not only focus on the fixed lowest price of the delivery.

“Consider during the procurement process that acquiring a tool that includes AI is not a one-time decision; testing the application over its lifespan is crucial.”

Considering the life-cycle management and its impact on procurement revealed to be a challenge. The earlier the focus on the maintenance of the solution and the ongoing management of the AI system, the better it is for the project delivery.

Success factors: Which guidelines were successfully implemented?

“Make use of innovative procurement processes to acquire AI systems - encourage collaboration between different bidders.”

It was important to rely on a team of suppliers for project delivery, rather than just one supplier. Partnering with three suppliers and asking them to deliver the project in collaboration ensures that all relevant skills were available and checks and balances were in place. Regarding AI delivery, one supplier developed the AI model and another supplier helped to test the model and ensured that it worked properly.

“Focus on developing a clear problem statement, rather than on detailing specifications of a solution.”

The requirements in the ITT focused on outcomes rather than the means of how to achieve those outcomes. This gave vendors the flexibility to select the technology that they found fit for purpose and ensure that the solution was innovative and effective.

“Work with a diverse, multidisciplinary team.”

The agency worked actively on upskilling internal teams and recruiting experts into the team where needed. This helped the agency to become a better customer for AI systems.

The delivery was supported through a close collaboration with the suppliers. During the project delivery the DVSA worked closely with delivery partners. Key to this was thinking as a single team and as partners, not contractors. At a practical level, this meant being open about the problems that needed to be solved, the challenges that different solutions may present and the costs of different options. This experience showed that openness brings real reward in getting value from the partnerships.

“Engage vendors early and frequently throughout the process.”

Extensive pre-market engagement helped to better target potential AI system providers. The DVSA hosted a supplier open day to explain the challenges that the agency faces to suppliers and gather initial ideas of how and with the help of which technologies to address these. After the initial tendering process, shortlisted suppliers were asked to present their approaches to the DVSA, which improved the ability to evaluate the different delivery approaches.

Below: Dubai

3 Case study United Arab Emirates

Dubai Electricity and Water Authority

Objective

To enable an efficient and comprehensive procurement process for digital and AI solutions, DEWA's top management had directed their team to demonstrate leadership on this topic. By identifying use-cases where the new procurement guidelines could be applied, DEWA's aim was to work on a pilot which could be then scaled across UAE and globally.

One of the use cases identified was the need for senior management at DEWA,to access reports and dashboards on a daily, weekly, monthly and quarterly basis to review strategic performance indicators. These dashboards and reports are available on different platforms and some of them take a long time to generate and prepare before they can be presented to top management. As a result, DEWA was looking for a faster and easier way to access the required data to make correct and timely decisions. A technology was needed that was capable of understanding management's enquiries, providing the right data in a convenient and timely way and learning from the enquiries made.

> “ The ability of machine learning to leverage a range of enterprise information and improve its interactions combined with the chatbot's ease of interaction proved to be an ideal means to meet the data access needs.

Why AI?

The use of AI to solve complex challenges was supported by the state's National AI Strategy, which seeks to position the UAE as an AI world leader by 2031. DEWA also has a vision to become a globally leading sustainable innovative cooperation, and its strategic objective is: “Enabling AI and digital technologies”. To achieve these goals, DEWA defined three main pillars for its AI adoption. The first is Rammas for You, which covers customer-facing services. The second is Rammas at Work, which seeks to augment the work environment with AI tools, and the third is and the first is Powered by Rammas, which adds AI to DEWA's core business assets.

In January 2017, DEWA launched the Rammas Virtual Agent, a chatbot that answers customers' enquiries and is powered by AI, as part of the Rammas for You pillar. Following the virtual agent's success, DEWA began considering using the same concept to meet management's data access needs. The ability of machine learning to leverage a range of enterprise information and improve its interactions combined with the chatbot's ease of interaction proved to be an ideal means to meet the data access needs.

Background

Dubai Electricity and Water Authority (DEWA) is a public utility founded on 1 January 1992, by a decree issued by the late Sheikh Maktoum bin Rashid Al Maktoum to merge Dubai Electricity Company and Dubai Water Department. DEWA's strategies and achievements are inspired and driven by the vision and directives of His Highness Sheikh Mohammed bin Rashid Al Maktoum, Vice President and Prime Minister of the UAE and Ruler of Dubai. Today, DEWA provides services to over 900,000 customers across Dubai.

DEWA was the 1st government organisation in the UAE to launch an online AI chatbot in 2017. The chatbot called Rammas communicates in both Arabic and English with customers and respond to their enquiries. AI helps DEWA's customers with services, such as the Smart Response service on DEWA's smart app and website. This allows early self-diagnosis of technical interruptions at home, reducing the necessary steps to deal with complaints and follow-ups.

DEWA conceptualised the AI procurement guidelines with the World Economic Forum and Dubai Future Foundation to further drive cooperation between the public and private sectors, and to enable governments and companies to make their procurement processes as efficient and transparent as possible by employing a multi-stakeholder approach. DEWA implemented a framework that allowed for feedback and finding best practices and standards to govern AI technologies procurement process.

Action

DEWA sent a request for proposal (RFP) to suppliers. Bidders had a month to respond, after which there was a window for bidders' questions and a bidder's conference to answer further questions.

The final evaluation of the solution proposals used seven criteria with different weights. Technical assessment and AI capability were the most important, and the proposed solutions were evaluated with a demonstration or evaluation of a prototype from each bidder's solution. DEWA also evaluated project governance, deliverables, business value, solution dependency and vendor background, and awarded the contract to the highest scoring proposal evaluated by the procurement committee, which comprises important stakeholders and AI specialists.

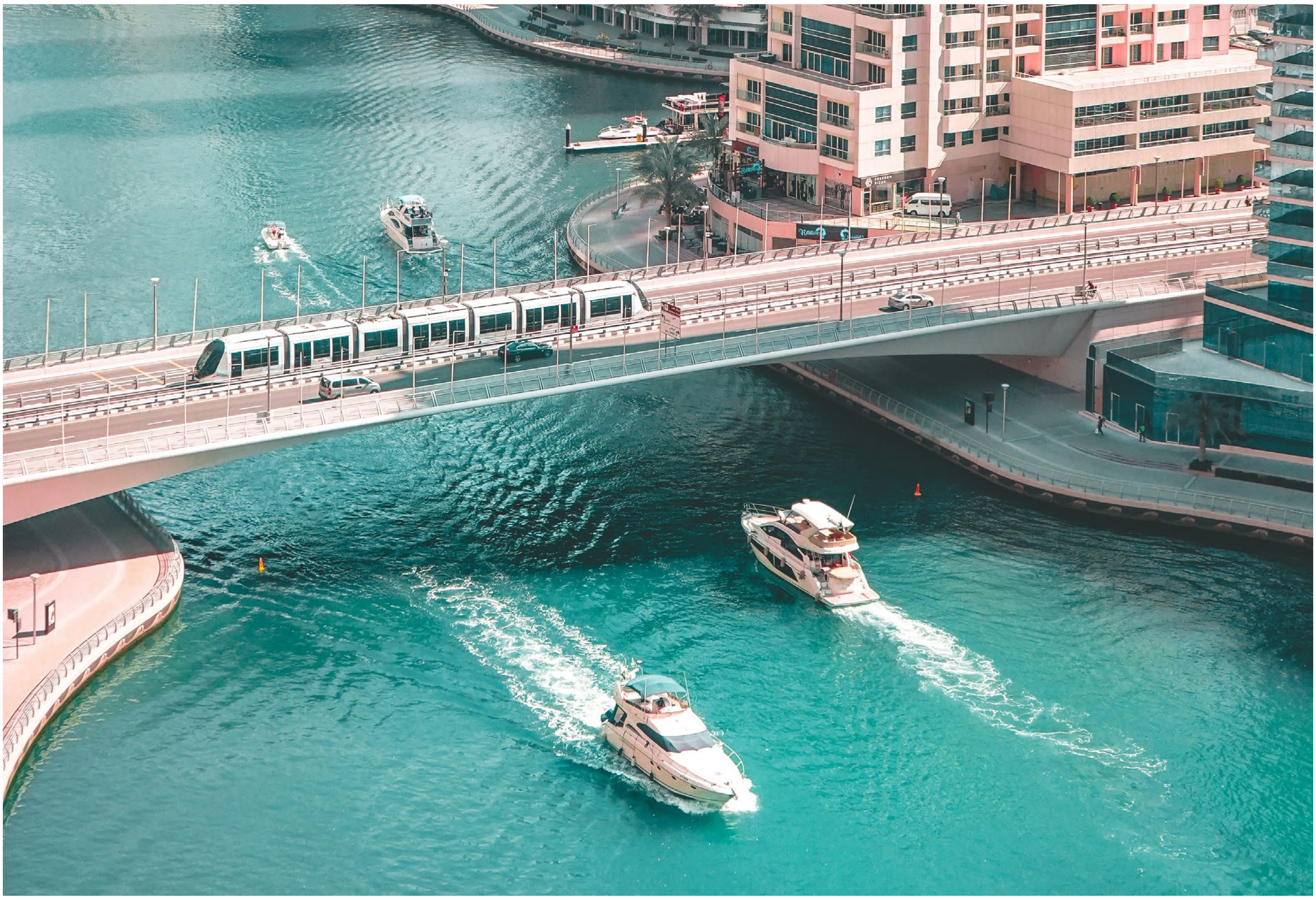
After this, the source code for the solution was shared with DEWA. This is an open source system and will be developed from scratch and hosted by MORO, a digital platform launched in 2018 to support the Dubai 10X initiative. MORO provides hosting and data storage services and cloud-based digital services management. The supplier contract took into account additional requirements, such as training DEWA employees to maintain and improve it to ensure continuity and the proper communication of knowledge, to enable DEWA to further expand its capabilities.

Ethical considerations

DEWA is committed to protecting customers' and stakeholder's data by adopting and complying with relevant UAE legislations and Dubai Government applicable regulations. This includes Federal Law No.1 for 2006 on Electronic Commerce & Transactions; Federal Legal Decree No. 5 for 2012 on combating cyber-crime, and the Regulatory Framework for stored values & Electronic Payment Systems (EPS Regulation), which regulates business offering electronic payment services.

DEWA also adheres to the Dubai Data Law, open data, shared data, data confidentiality and data sensitivity policies. DEWA also put in place internal measures to secure customer data. It drafted a contract that clearly stated the requirements to prevent sharing its information with any external parties; and that such data must always reside within DEWA's assets.

The solution works in tandem with multiple internal datasets related to strategic KPIs, employees' statistics, organisational data and sensitive information. The solution had to run on a private cloud within the UAE in adherence with the Dubai Data Law due to this sensitivity because it cannot be shared or processed externally. It was important that sensitive datasets remain protected at all times. To address this, the roles and responsibilities of each user were applied within the solution, and controlled by pre-defined access levels. There was considered to be no issue with data transparency or the ability to understand the AI model as the AI solution is only accessing data without any modification.



Lessons learned: Which guidelines were harder to implement?

“Make use of innovative procurement processes to acquire AI systems.”

The procurement process took five months from the business case initiation until the announcement of the preferred bidder. The team considered this would take too long. As a result, DEWA developed a new procurement track specific to AI in cooperation with the World Economic Forum and Dubai Future Foundation. This track was benchmarked by Dubai Future Foundation to apply similar techniques to specifically expedite the adoption of AI tools within DEWA.

The new procurement track consists of a set of key milestones including:

- Establish a senior AI Committee which includes champions from multiple departments and specialities to guarantee a 360-degree approach when evaluating AI RFPs and aligning them with AI Procurement Guidelines to ensure the adoption of the Framework, define an AI pre-approved supplier list, thus, enhance the overall AI procurement process and accelerate the adoption of AI technologies in DEWA.
- Create the DEWA AI Definition to have a clear description for AI-use cases within DEWA, avoid confusion with other technologies, and facilitate the overall process.
- Create AI RFP templates. Early market engagement will also be a key component of this new track, as the procurement team will constantly be on the search for new AI vendors via conferences and info sessions.

Success factors: Which guidelines were successfully implemented?

“Focus on developing a clear problem statement, rather than detailing specifications of a solution.”

DEWA implemented the first pilot for a virtual agent called Rammas, in 2016 and then launched the first version of the live solution in January 2017. Nine months later, the UAE AI Strategy was announced with a clear vision ‘to be an AI World Leader by 2031.’

The scope of the project was clear from the start as it was part of the AI roadmap initially. This made the process particularly efficient by leading to more relevant vendors’ responses and an increased probability of success.

“Aim to include your procurement within a strategy for AI adoption across government and learn from others.”

DEWA embedded AI in its strategy and developed a separate AI functional strategy that has been aligned and cascaded from the UAE National AI strategy. The functional AI strategy covers 6 main pillars, including AI in stakeholder happiness, AI in technology, AI in sustainability, AI in operations, AI investment, and enabling AI.

DEWA also responded immediately to the National AI Strategy by building a five-year roadmap to augment its work with AI tools. Moreover, DEWA is an active member of the Smart Dubai AI Advisory board and works closely with the Smart Dubai Office and other government entities for knowledge sharing and delivering new innovative services powered by AI to Dubai’s population.

For instance, the Rammas Virtual Agent content is integrated with Smart Dubai Office’s Virtual Agent, called Rashid, which is available on the Dubai Now smart application to ensure information availability and to maintain a seamless customer experience.

“Work with a diverse, multidisciplinary team.”

DEWA organised, in collaboration with Dubai Future Foundation, and World Economic Forum Fourth Industrial Revolution Centre, a four-day workshop in October 2019 about Artificial Intelligence (AI) Procurement guidelines.

This was part of DEWA’s efforts to position the UAE as a global leader in AI by 2031 in line with the UAE Strategy for Artificial Intelligence.

One of the main outcomes of this workshop was to form a senior AI committee within DEWA, which includes champions from the Contract and Procurement department, an AI Team, an Intellectual Property Team, the BRM Team, and the PMO Team. This committee is responsible for evaluating the AI RFPs and to align them with AI Procurement Guidelines, to ensure the adoption of the Framework, by defining an AI pre-approved supplier list, improving the overall AI procurement process, and accelerating the adoption of AI technologies in DEWA.

This ensured a comprehensive evaluation of the proposed solutions and a good understanding of the issues at play.

“Define if and how you will share data with the vendor(s) for the procurement initiative and the subsequent project.”

“Ensure that you have proper data governance mechanisms in place from the start of the procurement process.”

DEWA adheres to the Dubai Data Law, open data, shared data, data confidentiality and data sensitivity policies. Moreover, DEWA has internal measures to control data privacy. Customers’ data is not shared with any external parties and the data always resides within DEWA’s Assets.

DEWA’s security team is making sure that the data provided to the vendors is secured, encrypted and in compliance with Data Residency Law of UAE and DESC (Data Electronic Security Centre).

The Personal Identifier Information (PII) data was removed from the vendors’ dataset and the rest was encrypted. This gave the vendors access to the structure of the data, which is all that was needed to build a prototype.

4

Case study Kingdom of Bahrain

Information and eGovernment Authority

Objective

Decisions about advanced studies and career pathways in Bahrain have been traditionally based on strong cultural and social imperatives to pursue pure academic qualifications for traditional white-collar jobs, irrespective of whether there is labor market demand from those sectors. This social norm is compounded by the fact that there is no authoritative source of labor market intelligence on which prospective employees can base their study and career decisions. Together, these factors give rise to ill-informed decision-making, which has a detrimental impact on students, employers, and the government.

Therefore, the Labour Fund (Tamkeen) in collaboration with the Information and eGovernment Authority (iGA), and other government institutions, decided to develop an Employability Skills Portal (ESP) to serve as a repository of labor market information. This portal could be used by prospective employees to make informed career decisions and by educational institutions to tailor their programs to market demand. The portal needed a technology capable of cleaning and integrating data from multiple sources, finding correlation between the data and making prediction on the direction of various trends and indicators.

Why AI?

The use of AI was not a requirement at the start of the project. However, after discussion with different solution providers and evaluation of the first Proofs of Concept, it became clear that AI could add value to the proposed solution by using it for predictive analytics. In addition, the use of AI was in line with the vision of higher management and the Kingdom of Bahrain's leaders to support digital transformation and the use of modern technology.

> **After discussion with different solution providers and evaluation of the first Proofs of Concept, it became clear that AI could add value to the proposed solution by using it for predictive analytics.**

Background

The Information and eGovernment Authority (iGA) of Bahrain facilitates many public services related to the IT sector. It aims to achieve cyber security integration between the public sectors institutions, as well as to work on implementing the knowledge in order to support decision making, creativity and encouraging innovation in the areas of public services and institutions.

Action

As the portal would be based on the cloud, the project floated through an existing special procurement track for cloud technologies. This track accelerates the implementation of cloud projects by by-passing traditional tendering processes. In order to do that, this innovative procurement track offers access to dedicated funds for cloud technologies and a list of pre-approved vendors selected for their internal knowledge, links with global technology leaders and financial capabilities. The process started with a first, free of charge, Proof-of-Concept (POC), from different solution providers. These POCs were evaluated through an agile methodology until they reached an acceptable level of satisfaction by end users, the labor market, and internal users and iGA technical team. Each POC was then given a score based on both users' evaluation and a financial bid. Most weight was given to the ability to reach expected end results and user needs. The highest scoring vendor solution was chosen to move to the next phase; the development of a complete POC with costs covered by iGA. If the required level of satisfaction from the final POC was not met, iGA would select the next highest scoring vendor solution to move to the second phase until the required level of satisfaction was reached and the contract was awarded. This iterating phase took about two months to complete.

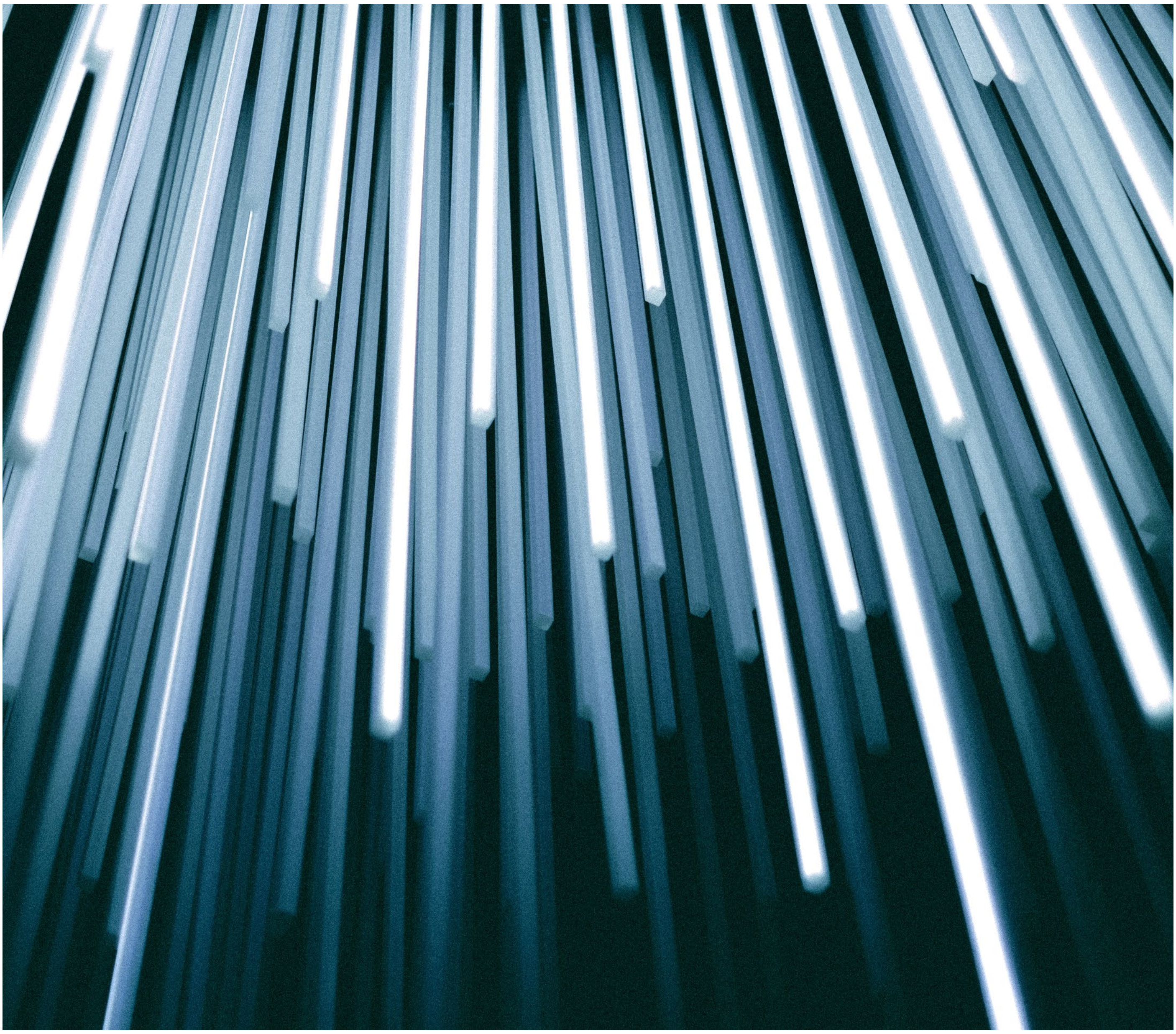
The solution was agreed to be fully owned by iGA and its internal technical team was involved from the start in the implementation process to ensure a proper handover of the solution. iGA technical and management team also made sure to benefit from the bidders' knowledge through weekly meetings and close collaboration to better understand the implication and use of AI.

Ethical considerations

The Data Protection Law of Bahrain, which regulates the use of personal data, was taken into consideration for the project and vendors had to comply with it. However, the project also involved other types of data not covered by the regulations. The use of various datasets from different government entities was an important issue because of the lack of regulations and governance for data sharing between organizations and the lack of governance for non-personal data. Hence, a task force leaded by iGA and top management from each involved organization was created. The role of this task force was, in part, to serve as a governance body for data sharing and also to gain an in depth understanding of each dataset and the biases that could emerge when using AI. Indeed, the best way to gain insights on the nature of each dataset and their potential bias was to partner with the providers of these datasets. iGA also appointed an external legal consultant to conduct an impact assessment of the use of data before starting the project. The goal was to understand if the way each dataset would be used could create legal or ethical issues.

Concerning data sharing with the vendors for the POCs development, the vendors had access to the entire population to train their models, but synthetic data was used to mask personal information. The synthetic data was generated in such a way that the real aggregate results were preserved. In addition, the vendors could only access the data through temporary iGA internal accounts.

The AI model explainability was addressed by requiring the successful vendor to provide a non-technical description of the model that would be available to internal users.



Lessons learned: Which guidelines were harder to implement?

“Create the conditions for a level and fair playing field among AI solution providers.”

The accelerated cloud technology procurement track being a new process, the list of pre-approved vendors was not fully developed at the time of the project. Work is being done to expand this list and give access to new innovative vendors.

“Make use of innovative procurement processes to acquire AI systems.”

The introduction of payment for the development of the second POC was a new concept that slowed the process as it was hard to get approvals. Moving forward, instead of requiring approval for each new payment, the accelerated cloud technology procurement track will include lump-sum funds that can be allocated as needed for each procurement project.

Success factors: Which guidelines were successfully implemented?

“Focus on developing a clear problem statement, rather than on detailing specifications of a solution.”

The project didn’t start with AI in mind. The need for a specific outcome was defined and the technical evaluation of the vendors’ solution was focused on their capacity to meet the desired outcome. Hence, the project was open to a variety of technical solutions and was able to select the most appropriate technology.

“Conduct an initial AI risk and impact assessment even before starting the procurement process, ensure that your interim findings inform the RFP, and revisit the assessment at decision points.”

An external consultant was mandated to evaluate the potential impacts of the use of AI on the different datasets. Potential biases were identified as well as the mitigation strategies.

“Conduct a review of relevant legislation, rights, administrative rules and other relevant norms that govern the types of data and kinds of applications in scope for the project.”

“Ensure that you have proper data-governance mechanisms in place from the start of the procurement process.”

Relevant regulations were identified and communicated to the vendors. In addition, blind spots within the current regulations were identified and strategies were put in place to address them. A government task force was formed to identify best practices and establish consensus on the use, processing and transfer of non-regulated data.

“Highlight the technical and ethical limitations of using the data to avoid issues such as bias.”

A government task force comprised of top management from each organization where data would be collected was created. Hence, the vendors and iGA team were able to meet with the data providers and truly understand potential biases and limitation to the quality of each datasets in order to avoid misleading results. Vendors were then able to adapt their model accordingly and address these shortcomings.

“Ask the AI provider for knowledge transfer and training to be part of the engagement.”

iGA internal technical team was involved from the start in the implementation process to ensure a proper handover of the solution. iGA technical and management team also made sure to benefit from the AI providers’ knowledge through weekly meetings and close collaboration to better understand the implication and use of AI.

5

Case study Splunk Inc.

Key considerations for successful adoption of AI as an added capability/functionality with an existing supplier and a system already in use

There are multiple ways of procuring and adopting AI technologies; they can be built from scratch, added as capabilities to commercial off-the-self (COTS) systems or acquired directly as a service (SaaS). Often solutions require a mix of these approaches to be successfully adopted. For most operational organizations AI capabilities are added iteratively to an existing solution or procured via an existing supplier as an added functionality to a product or service. When adopting AI as part of an existing platform contract without going through an independent AI procurement process, some guidelines are more relevant than others.

Three important factors, highlighted in the guidelines, form the basis of success for public-sector agencies adding AI-capabilities to systems already in use. These have emerged from Splunk's experience supporting and working collaboratively with public-sector entities:

Key guideline for AI as an added capability/functionality	Key factor to consider to successfully implement the guidelines
"Define the public benefit of using AI while assessing risks."	**End users' background** When considering the benefits that can be realized with an AI system, understanding the end-user audience is of great importance. The end-user's understanding of pertinent mathematical principles (such as probability) and how they are likely to interpret and apply the output of the AI system should be considered. This will help inform the type and granularity of outputs (e.g. visual charts, key metrics etc.) that should be selected, how fast new techniques can be adopted and/or accepted and what cautions, if any, are desirable for the particular use case.
"Articulate the technical and administrative feasibility of accessing relevant data."	**Understanding data assets** Finding and understanding what data an organization holds and how it may be accessed, combined and processed in accordance with the law and organizational norms will help you determine project scope – what can be achieved with the data and with what controls. According to recent research, 97% of public-sector agencies agree that they must improve their ability to ingest, index and cross-correlate disparate data sets to optimize public policy outcomes.
"Highlight the technical and ethical limitations of intended uses of data to avoid issues such as historical data bias."	**Data literacy** AI technologies can be complex and therefore, to be successful in the identification of technical and ethical limitations, it is critical that an organization's leadership and operations team be "data literate". This does not mean each team member must become a data scientist, but they should understand the underlying mathematical principles (i.e. probability, accuracy, sampling etc.) and gain an appreciation of the different benefits and limitations of the main ML techniques. Innovation and education go hand in hand. Without a proper data and knowledge foundation, users will not be able to capitalize on the advances in automation and decision-making capability provided by AI.

Acknowledgements

The World Economic Forum's Unlocking Public Sector Artificial Intelligence project, in collaboration with the Government of the United Kingdom, Deloitte Consulting and Splunk is a global, multistakeholder and cross-disciplinary initiative intended to help shape the public sector's adoption of AI, and emerging technologies in general, around the world. The project has engaged leaders from private companies, governments, civil society organizations and academia to understand public-sector procurement of AI technology, identify challenges and define principles to guide responsible and ethical procurement. The opinions expressed herein may not correspond with the opinions of all members and organizations involved in the project.

Lead authors:

Sabine Gerdon
Artificial Intelligence and Machine Learning Fellow, World Economic Forum, Seconded from the Office for Artificial Intelligence, Government of the United Kingdom

Eddan Katz
Project Lead, World Economic Forum

Emilie LeGrand
McGill University Integrated Management Student Fellow

Gordon Morrison
Director of EMEA Government Affairs, Splunk Inc.

Julián Torres Santeli
Artificial Intelligence and Machine Learning Fellow, World Economic Forum, Seconded from Deloitte Canada's AI practice

We would like to thank our Unlocking Public-Sector AI project community as well as the following contributors for their insights:

Rashid Alahmedi
Senior Specialist Technolgy and Solutions, Dubai Electricity and Water Authority

Greg Ainslie-Malik
Machine Learning Architect, Splunk Inc.

Jesus Alvarez-Pinera
Head of Data, Food Standards Agency

Shelby Austin
Managing Partner, Growth and Investments and Omnia AI, Deloitte

Yousef Al-Barkawie
Partner, Analytics and Cognitive Middle East Leader, Deloitte

Neil Barlow
Head of Vehicle Policy and Engineering, Driver and Vehicle Standards Agency

Kathy Baxter
Architect, Ethical AI Practice, Salesforce

Lorena Cano
Digital Trade Fellow, World Economic Forum from Inter-American Development Bank

Ashley Casovan
Executive Director, AI Global

Michael Costigan
Artificial Intelligence and Machine Learning Fellow, World Economic Forum from Salesforce

Sue Daley
Associate Director, techUK

Nihar Dalmia
Government and Public Sector AI leader for Deloitte Canada, Deloitte

Gourav Dhiman
Business Development Manager, XLPAT

Cosmina Dorobantu
Deputy Director of Public Policy Programme, The Alan Turing Institute

Leslie Harper
Senior Sector Specialist, Inter-American Development Bank

James Hodge
Chief Technical Adviser, Splunk Inc.

Hamad Karam
Senior Specialist Artificial Intelligence, Dubai Electricity and Water Authority

Andrew Kim
Head of AI Policy, Google Cloud

Steven Knight
AI Lead, Food Standards Agency

Benjamin Leich
Economic Adviser, Better Regulation Executive

Katherine Mayes
Programme Manager, techUK

Maha Mofeez
Chief Corporate Officer,
Bahrain Economic Development Board

Valesca Molinari
Automotive and Autonomous Mobility Fellow,
World Economic Forum from Baker McKenzie

Mariam Al Muhairi
Head, Centre for the Fourth Industrial Revolution
United Arab Emirates

Khalid Al Mutawa
Director, Bahrain Information
and eGovernment Authority

Brandie Nonnecke
Founding Director, CITRIS Policy Lab

Arwa Al Qassim
AI Lead, Centre for the Fourth Industrial Revolution
United Arab Emirates

Ana Rollan
Artificial Intelligence and Machine Learning Fellow,
World Economic Forum from BBVA

Nada Al-Saeed
Data Policy Fellow, World Economic Forum from
Bahrain Economic Development Board

Komal Sharma Talwar
Director, XLPAT and TT Consultants

Leonard Stein
Senior Strategic Adviser, Splunk Inc.

Jitin Talwar
Founder, XLPAT and TT Consultants

Sandeep Singh Kohli
Co-founder, XLPAT

Ahmad Al Tawallbeh
Specialist Artificial Intelligence, Dubai Electricity
and Water Authority

Abbey Thornhill
Assistant Economist, Better Regulation Executive

Adrian Weller
Programme Director for AI, The Alan Turing Institute

Mark Woods
Director, Technology and Innovation, Splunk Inc.

Tim Woodbury
Director of State and Local Government Affairs,
Splunk Inc.

Thank you also to the teams in the UK from the Defence Science and Technology Laboratory, the Department for Transport, the Home Office Accelerated Capability Environment and local governments that supported the user testing and piloting. The steering and working group from the Department of Digital, Culture, Media and Sport, the Government Digital Service, the Cabinet Office, the Crown Commercial Service and the Centre for Data Ethics and Innovation has been instrumental to progressing this work, in particular:

Sue Bateman
Deputy Director for Policy and Innovation,
Government Digital Service

Oliver Buckley
Executive Director, Centre for Data Ethics and
Innovation

Stephen Hennigan
Deputy Head of Office for Artificial Intelligence,
United Kingdom Government

Sana Khareghani
Head of Office for Artificial Intelligence, United
Kingdom Government

Thank you to everyone who contributed through interviews, workshops and discussions in the last 18 months in Dalian, Dubai, London, Manama, San Francisco, Tianjin, Toronto and Washington DC.

Endnotes

1. https://www.gartner.com/en/information-technology/glossary/open-source
2. https://searchdatacenter.techtarget.com/definition/COTS-MOTS-GOTS-and-NOTS
3. https://www.gartner.com/en/information-technology/glossary/infrastructure-as-a-service-iaas
4. https://www.gartner.com/en/information-technology/glossary/platform-as-a-service-paas
5. https://www.gartner.com/en/information-technology/glossary/software-as-a-service-saas
6. https://www.sfia-online.org/en
7. Factsheets: Increasing Trust in AI Services through Supplier's Declarations of Conformity. Matthew Arnold, Rachel K. E. Bellamy, Michael Hind, Stephanie House, Sameep Mehta, Aleksandra Mojsilovic, Ravi Nair, Karthikeyan Natesan Ramamurthy, Darrell Reimer, Alexandra Olteanu, David Piorkowski, Jason Tsay, Kush R. Varshney. https://arxiv.org/abs/1808.07261
8. API stands for application programming interface. AnAPIis a software intermediary that allows two applications to talk to each other.
9. https://www.gov.uk/guidance/digital-outcomes-and-specialists-buyers-guide
10. Unsupervised learning was used given the team did not have labelled data.

WORLD ECONOMIC FORUM

COMMITTED TO
IMPROVING THE STATE
OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org