

Responsible AI for net zero

An Australia and India collaborative approach towards practice, governance, and ethics in energy future

SWINBURNE UNIVERSITY OF TECHNOLOGY

Australian Government
Department of Foreign Affairs and Trade

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

INDIAN INSTITUTE OF TECHNOLOGY PALAKKAD

Responsible AI for net zero

An Australia and India collaborative approach towards practice, governance, and ethics in energy future

Australian Government
Department of Foreign Affairs and Trade

Project partners

Supporting partners

Acknowledgement

This project is supported by the Australian Government Department of Foreign Affairs and Trade (DFAT) under the Australia-India Cyber and Critical Technology Partnership (AICCTP).

Swinburne University of Technology and its research partners remain independent in the content of the research on this paper and in all editorial judgements.

Acknowledgement of Country

In the spirit of reconciliation, we acknowledge the Traditional Custodians of Country throughout Australia and their connections to land, sea and community. We pay our respect to their Elders past and present and extend that respect to all Aboriginal and Torres Strait Islander peoples today.

Project team

Australia project team members

Swinburne University of Technology

Professor Prem Prakash Jayaraman

Professor Anthony McCosker

Dr Abhik Banerjee

Professor Mehdi Seyedmahmoudian

Dr Yong-Bin Kang

Professor Alex Stojcevski

Dr Gokul Thirunavukkarasu

Dr Kafeel Ahmed

Dr Robert Lundberg

India project team members

Indian Institute of Technology Palakkad

Professor Vijendran Venkoparao

Associate Professor Narayanan C Krishnan

Indian Institute of Technology Madras

Professor Balaraman Ravindran

Associate Professor Satyanarayanan Seshadri

Dr Ramkishore Sankaralingam

Supporting partner organisations

Nasscom

Mr Praveen Mokkapati

Indian Institute of Technology Gandhinagar

Associate Professor Naran Pindoriya

Publisher

Swinburne University of Technology
Hawthorn Victoria 3122 Australia
swin.edu.au
First published 2024

DOI: doi.org/10.25916/sut.26866951

Copyright © 2024 the authors.

This is an open access work distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY NC 4.0) license. See creativecommons.org/licenses/by-nc/4.0/

Suggested citation: Prem Prakash Jayaraman, Anthony McCosker, Abhik Banerjee, Yong-Bin Kang, Gokul Thirunavukkarasu, Kafeel Ahmed, Robert Lundberg, Mehdi Seyedmahmoudian, Alex Stojcevski, Vijendran Venkoparao, Narayanan C Krishnan, Satyanarayanan Seshadri, Balaraman Ravindran (2024), Responsible AI for Net Zero: An Australia and India Collaborative Approach Towards Practice, Governance and Ethics in Energy Future, Swinburne University of Technology. doi.org/10.25916/sut.26866951

Executive summary

Artificial intelligence (AI) is poised to reshape all parts of the global energy sector. While this critical set of evolving technologies offers vast potential in achieving energy efficiency and sustainable net zero goals, it also presents ethical and security risks.

This project – *Responsible AI for net zero: An Australia and India Collaborative Approach towards Practice, Governance and Ethics in Energy Future* – has been funded by the Australian Government Department of Foreign Affairs and Trade (DFAT), under the Australia-India Cyber and Critical Technology Partnership (AICCTP) Round 3. It aimed to:

- Establish current knowledge and ethical principles relevant to AI adoption in energy sector in Australia, India and Indo-Pacific region,
- Identify key ethical issues shaping the use of AI in developing smart energy systems to achieve 'net zero', and
- Recommend guidelines for translating ethical principles into AI energy governance practice.

Our research, informed by interactive research workshops with energy sector stakeholders in Australia and India (Chennai, Gandhinagar and Melbourne) and a comprehensive review of current research, underscores the transformative potential and inherent challenges of AI in the journey towards net zero.

Drawing on diverse perspectives from across the sector in these two national contexts this report will help policy makers and stakeholders from across the energy sector chart a pathway toward the implementation of responsible AI principles.

AI's promise and potential

Stakeholders in the energy sector are becoming well-versed in AI applications to achieve net zero. AI is primarily viewed as a powerful tool for prediction and forecasting, such as predictive maintenance of grid infrastructure. AI technologies also hold the promise to revolutionise smart energy systems, which is crucial for sustainable energy practices and achieving net zero goals.

Managing the risks of AI

Responsible AI is about contributing to positive outcomes while mitigating risks. Effective AI integration requires the energy sector's systems, processes, and technologies to be AI-ready, particularly in improving data gathering and data and AI governance. Data quality is crucial for ensuring AI makes accurate, privacy-preserving and unbiased decisions. Despite the risks, there is increasing interest in AI's transformative potential for the energy sector.

Essential responsible AI principles

Stakeholders understood that mistakes made by AI should have clear human accountability, ensuring that responsibility is attributable. AI must respond to local conditions while maintaining consistency across contexts. Moreover, AI usage decisions should involve diverse stakeholder input. Despite differences in the energy sectors in Australia and India, participants understood that implementing a shared set of responsible AI principles is key to stakeholder acceptance, with ethical, equitable, and sustainable AI deployment as priorities for both nations.

Need for a clear roadmap to embed responsible AI

A well-defined roadmap is crucial for identifying and prioritising key aspects of successful AI adoption in the energy sector. Data and detailed analysis are essential for developing regulations that respond to local conditions. Realising responsible AI across the sector requires multi-stakeholder collaboration, including government, industry, academia, energy regulators, and startups. Our roadmap serves as a foundational step towards integrating responsible AI principles into the energy sector, aiming for a resilient and sustainable energy future supporting net zero. Further work and collaboration are necessary to achieve this vision.

Contents

Acknowledgement 2
Project team 2

Executive summary 3

Introduction 5

Research methodology 6
Desktop research and analysis 7
Interactive research workshops in Australia and India 7
Workshop data analysis 9

Responsible AI for net zero – A global research landscape 10
Applications in the energy sector for achieving net zero 11
Considerations for responsible use of AI 12

Research workshops – An Australia India collective 13
Workshop participants 13
Workshop session themes 14

Foundations for establishing responsible AI for net zero guidelines 15
The role of AI in supporting the transition to net zero 15
Responsible AI: risks, challenges and opportunities 17
Best practices and considerations for governing AI 20
Data readiness 20
Accountability 20
Regulation 21

Guidelines for translating ethical principles into practical responsible AI energy governance 22

Summary 24

References 26

Glossary of terms 28

Appendix A – Systematic literature review 29

Appendix B – Research Workshops in Australia and India 30

Appendix C – Research workshop worksheets 33

Introduction

The transition to net zero relies on development of digital tools for the energy sector, with Artificial Intelligence (AI) being a key factor in achieving 'net zero' targets (2050 – Australia and 2070 – India)[1].

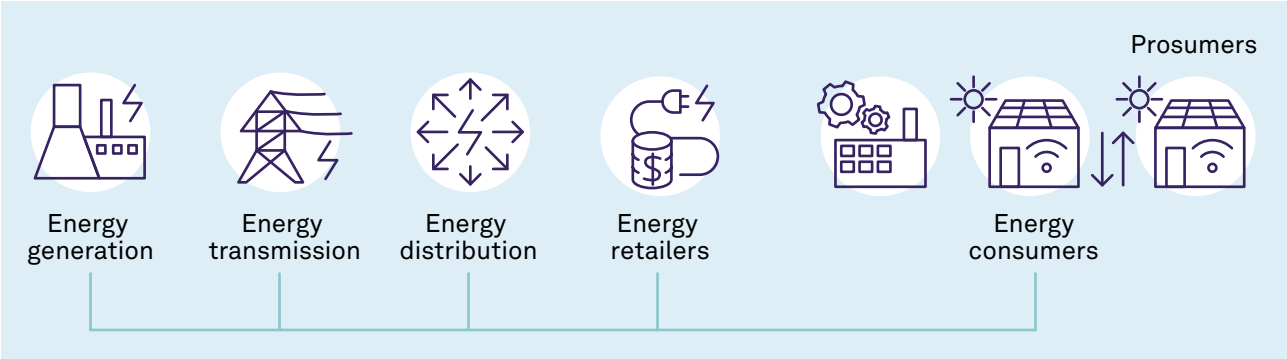
The use of AI puts significant pressure on ethical values and fundamental rights that drive our societies. The ethical concerns arise not only from privacy violations regarding collection and use of sensitive consumer/prosumer data, but also from the imperative to safeguard energy equity and affordability while maintaining sustainability and resilience to emerging security threats to energy infrastructure.

Through a rigorous stakeholder engagement approach across Australia and India this project brought together utility providers, regulators, local government, and businesses to develop a collective nuanced understanding of the challenges and opportunities associated with integrating AI into energy systems towards achieving net zero.

This report presents the findings of this project and the research conducted into identifying the challenges and opportunities as well as relevant guidelines in the responsible and ethical deployment and adoption of AI in the energy sector. This project has contributed to the global discourse on responsible AI development and deployment in the energy sector.

We hope that the insights and recommendations presented in this report will provide valuable knowledge for policymakers, industry leaders, researchers, and stakeholders across the energy value chain and others committed to establishing responsible AI practices that support the transition to a clean and sustainable energy future for achieving net zero goals.

Figure 1: *Energy value chain*

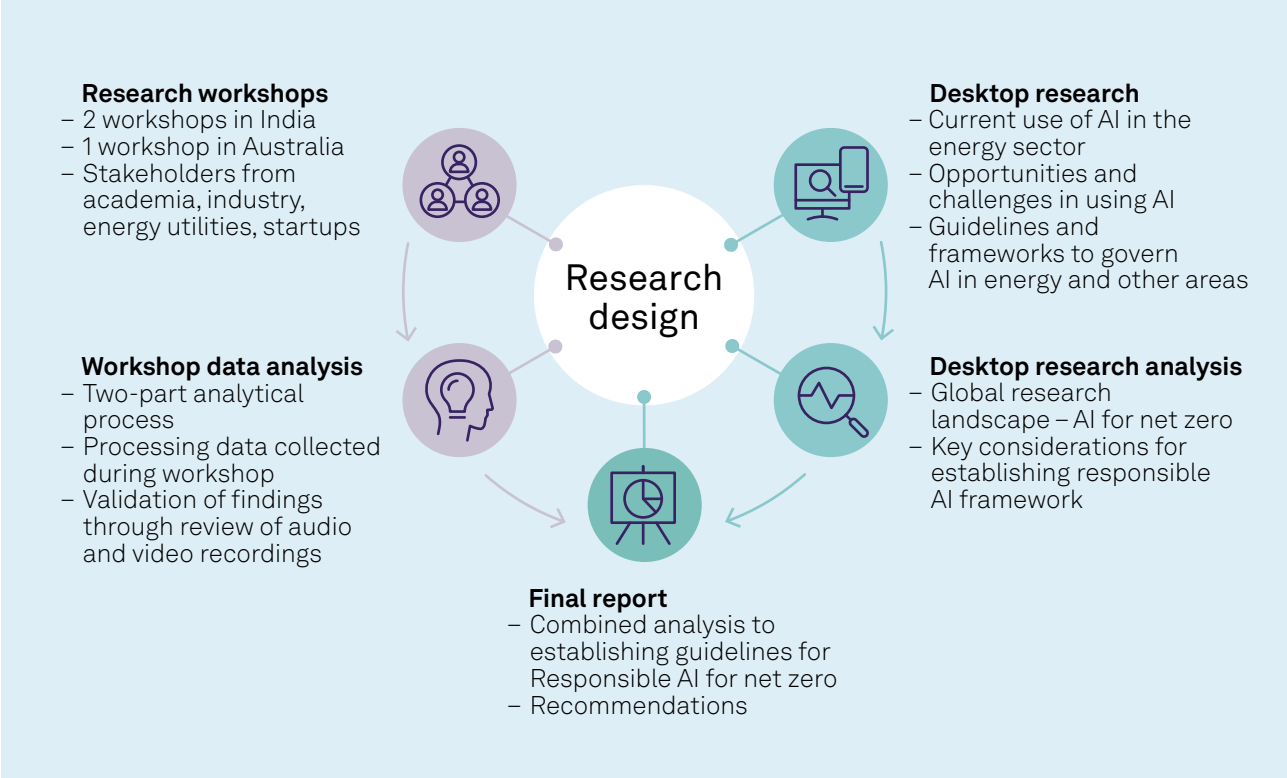


Research methodology

This project employed a combination of desktop research and interactive qualitative research workshops to capture data from a range of stakeholders.

Participant consent protocols were approved by Swinburne's ethics committee (20237513-17124). Our research design is presented in Figure 2 below and briefly elaborated further in this section.

Figure 2: *Research design adopted for the project*



Desktop research and analysis

The objective of our desktop research was to conduct a systematic literature study to and identify the use cases of AI within the energy sector specifically around the development of Smart Energy Systems (SESs) which are fundamental in reaching net zero. We conducted searches on Google Scholar, IEEE Xplore, and ACM Digital Library, focusing on academic papers and industry reports published between 2020 and 2023. Based on our preliminary review, we selected 48 papers (see Figure 12 in Appendix A) and reports that closely aligned with the project's objectives.

This desktop research analysis aimed to:

- Explore the local (Australia and India) and global landscape of AI and Responsible AI in the development of SESs for achieving net zero in the energy sector.
- Guide the development of the interactive research workshop.
- Inform the current state and future potential of AI and Responsible AI in the energy sector and identify the key areas for further exploration and development.

Interactive research workshops in Australia and India

We conducted three full-day interactive research workshops: two in India (Chennai and Gandhinagar) and one in Melbourne. These workshops aimed to engage a range of stakeholders including academics, researchers, energy sector professionals, and other relevant stakeholders across the energy value chain to discuss their understanding, knowledge and related challenges and opportunities of responsible use of AI in the energy sector. Specifically, the workshops aimed to:

- Engage diverse stakeholders from both countries to discuss their knowledge of Responsible AI.
- Identify and discuss use cases for AI in the energy sector.
- Explore the knowledge of energy sector stakeholders about AI and specifically Responsible AI.
- Gauge the potential of AI as a tool to support net zero and importance of using it responsibly.
- Speculate on the potential implications in responsible use of AI in the energy sector.
- Understand key guidelines and stakeholders that need to be engaged to ensure AI is implemented responsibly.

To achieve this aim, each of the workshops in Australia and India followed similar formats. The workshops included three sessions over the course of the day. In India, speakers from industry, academia, and the project team gave short presentations at the beginning of each session. These presentations contextualised Responsible AI, AI in the energy sector, and why we are interested in the responsible use of AI to support our transition to net zero.

Next, the participants broke off into small breakout sessions (groups) for facilitated discussions. The composition of the groups was decided by the project team prior to the workshop based on registered participants. The project team considered diversity and sufficient representation from stakeholder category (academia, industry, utility providers, government) when composing the groups.

To focus these discussions, the team used posters with questions representing the corresponding session and space for participants to note their responses using Post-it notes. In addition to this written record, each conversation was audio and video recorded, and photographs were taken of the posters and Post-it notes. Further, each group also had one of the project team members as part of the group to help facilitate the conversations. Finally, a representative from each small group summarised the key points of discussions to all workshop participants.

Figure 3: *A snapshot of interactive research workshops in Australia (Melbourne) and India (Chennai and Gandhinagar)*







Workshop data analysis

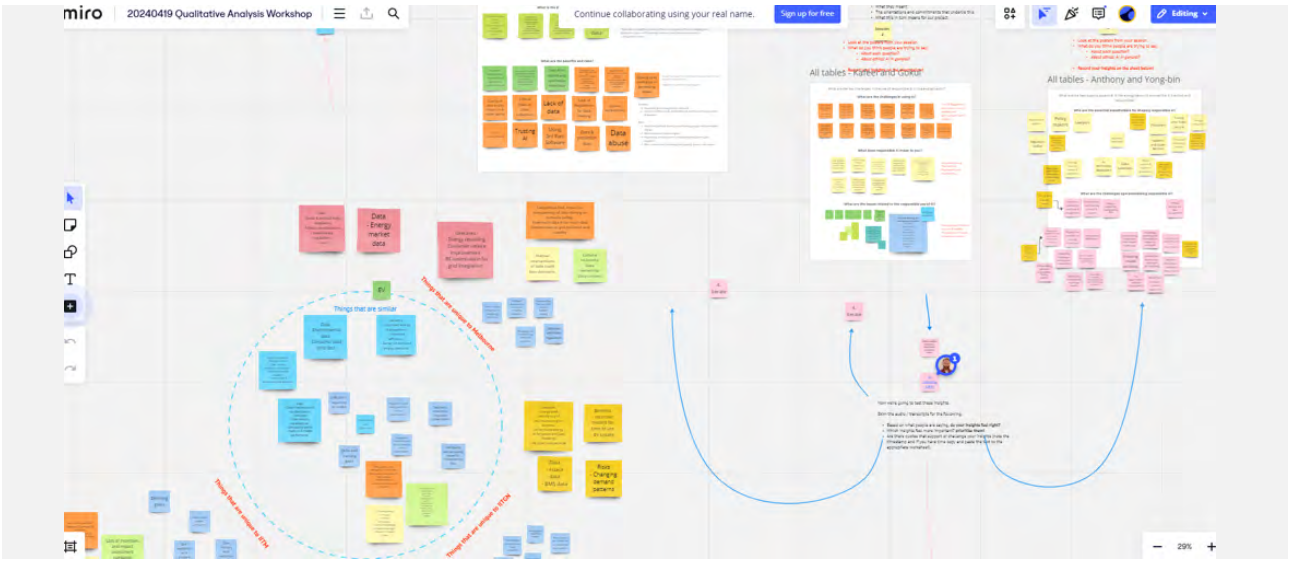
This activity involved a thorough examination of data collected from all the workshops. The analysis aimed to aggregate and analyse both individual workshops as well as collective feedback and discussions made through qualitative analysis. We employed a two-part qualitative analysis process:

1. **Thematic analysis:** Analysed the data collected from the workshops, organised by the various themes of the questions discussed. Specifically
   - Reviewed posters and participant responses noted on Post-it notes.
   - Summarised key ideas/themes that highlighted from group conversations.
   - Used visual whiteboard software Miro to record and organise key ideas.
   - Compared and contrasted ideas from different workshops to identify similarities and differences.
2. **Validation findings through recordings:** Validated thematic analysis results by revisiting audio and video recordings.
   - Reflected on key ideas/themes developed during the thematic analysis.
   - Compared and contrasted themes with participants' verbal discussions.
   - Captured key quotes to support the discussion.

Figure 4: *Project team sifting through data collected from workshop*



Figure 5: *Project team using Miro board for organising and analysing the workshop participant responses*



Responsible AI for net zero – A global research landscape

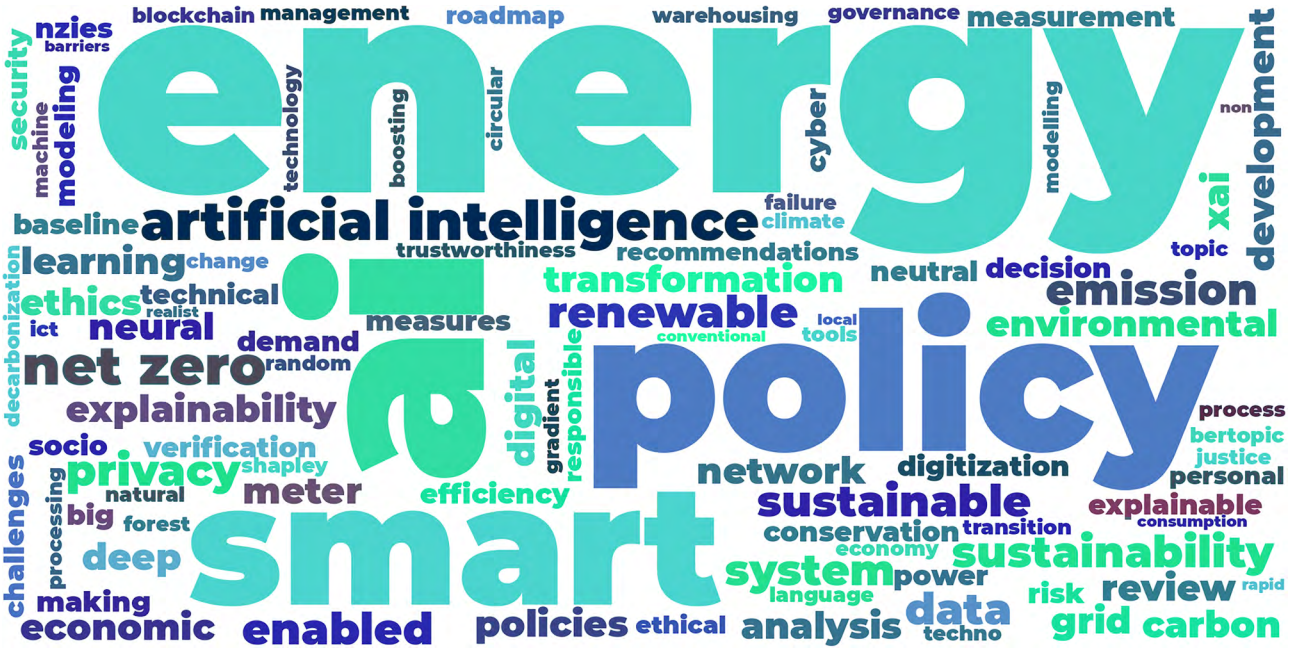
In this section we present our analysis of the desktop literature.

The word cloud in Figure 6 highlights the most frequently occurring keywords in our surveyed literature (see Figure 12 in Appendix A). These keywords present the key themes and focal points of discussion in the current literature across the areas of AI, responsible AI, net zero, and energy.

Some of the key findings include:

- **Energy and policy:** The prominence of "energy" and "policy" underscores the emphasis on energy management and regulatory frameworks.
- **AI integration:** Frequent mentions of "artificial intelligence" and "smart" technologies reflect the increasing role of AI and digital technologies in the energy sector.
- **Sustainability and renewable energy:** Keywords such as "renewable," "sustainable," and "net zero" highlight the shift towards environmentally friendly energy systems.
- **Data and analysis:** The importance of "data," "analysis," and "system" emphasises the need for data-driven decision-making.
- **Ethics and privacy:** The inclusion of "ethics" and "privacy" indicates growing concerns about ethical considerations and data privacy in smart energy systems.

Figure 6: *A word cloud of the top keywords from research spectrum in smart energy systems*



Applications in the energy sector for achieving net zero

We categorised AI applications used in the current energy systems based on four major areas of the energy supply chain:

- **Energy generation:** Optimise and increase renewable energy use and maximise efficiency of energy generation.
- **Energy transmission:** Enhance grid stability and reliability through predictive analytics, foreseeing and mitigating potential disruptions.
- **Energy distribution:** Dynamically balance supply and demand, enabling efficient energy flow, minimise energy waste.
- **Energy consumption:** Accelerate the development of intelligent energy management systems for consumers by learning and adapting to energy consumption patterns, controlling energy consumption, and reducing energy losses.

Table 1 provides a summary of AI applications and how they help support net zero transition across the above four major areas.

Table 1: AI applications, their objectives, and their impacts associated with net zero goals
(1: Increasing renewable energy adoption, 2: Enhancing energy efficiency, 3: Advancing grid resilience)

Energy systems	Application	Objective	Net zero impact		
			1	2	3
Energy generation	DERs (Distributed Energy Resources) optimisation[2]	To plan energy generation based on available resources.	✓	✓	✓
	Load forecasting[3]	To forecast electricity demand learning based framework based on historical data.	✓	✓	
	Generation forecasting[4]	To predict energy generation to facilitate energy management.		✓	✓
	Renewable integration Planning[5]	To expand the adoption of renewable energy sources.	✓		
Energy transmission	Power flow optimisation[6]	To provide voltage stability by optimising power flow of transmission network.		✓	✓
	Grid stabilisation[7]	To balance the generation and demand of energy in power system.		✓	✓
	Fault detection and diagnosis[8]	To detect and classify transmission line faults.			✓
Energy distribution	System improvement planning[9]	To reduce active power losses by optimising shunt capacitor installation locations.			✓
	Network loss reduction[10]	To tune PID controllers for controlling microgrid grid voltage.		✓	✓
	Fault detection and diagnosis[11]	To diagnose energy network fault.		✓	✓
	Demand response management[12]	To forecast and generate demand response messages for load changes and optimising the demand curve.		✓	
Energy consumption	Home energy management[13]	To optimise energy consumption in response to changes in supply, price, and demand.	✓	✓	
	Smart buildings efficiency[14]	To regulate the temperature of buildings.		✓	
	Fault detection[15]	To detect and diagnose sensor faults in building energy systems.		✓	✓
	Theft detection[16]	To distinguish between normal electricity consumption and suspected pattern.		✓	✓

Considerations for responsible use of AI

As noted from the literature, AI is recognised as a key pillar of future sustainable energy systems and offers significant advantages in achieving our net zero goals. However, the use of AI also raises key questions about its responsible use. In the energy sector, a responsible AI paradigm can help guide the use of AI ethically, equitably, and sustainably. Table 2 presents seven key dimensions of responsible AI principles that need to be considered for development and deployment of responsible AI-based energy systems.

These dimensions have been identified from a synthesis of our desktop research[17] [18] [19] [20] on ethical AI principles and responsible AI practices that have been highlighted by researchers as important for future research topics. This synthesis reflects the evolving understanding of how to apply AI technologies responsibly in the energy sector.

Table 2: Dimensions of responsible AI in smart energy systems

Dimension	Core concept	Key considerations
Data	Promote privacy, data protection, bias, fairness, and security. Responsible data practices are vital for developing and using AI systems ethically and socially.	Ethical data use, preventing bias, data security
Algorithm	AI technologies must support efficient energy management while being transparent, explainable, accountable, and responsible. This fosters trust, fairness, and societal well-being.	Algorithmic transparency, explainable AI, accountability
Human-centric	AI should prioritise human-centric values and well-being, including user consent, autonomy, human impact, and safety.	User consent, autonomy, human impact and safety
Social and societal	Ensure AI's positive impact on society by adhering to principles of fairness, non-discrimination, social responsibility, and beneficence, establishing the equitable and ethical application of AI.	Fairness, non-discrimination, social responsibility
Governance and regulatory	Guide that AI applications align with legal frameworks and cultural expectations, including regulatory compliance, adherence to laws, and global and cultural norms.	Regulatory compliance, adherence to laws, cultural norms
Environmental	Promote sustainability and minimises the environmental footprint of AI technologies.	Sustainability, environmental stewardship
Economic and labour	Focus on the economic impacts and labour considerations of AI applications, ensuring they contribute to economic growth and address labour market dynamics responsibly.	Economic growth, labour market dynamics

Research workshops – An Australia India collective

Workshop participants

The workshops attracted a diverse group of participants from various sectors and locations across Australia and India.

The Figure 7 illustrates the distribution of participants by organisation type. Academics made up the majority, accounting for 51.7 % of the attendees. This was followed by equal representation from the Small medium-sized enterprises (SMEs)/Startups and Industry sectors, each contributing 16.7 %. Participants from Electric Utility represented 11.7 %, while Government representatives made up 3.3 %. The workshop had 83 % male and 17 % female participants, which further highlighting the need for increased gender diversity in the energy sector which has been identified as a challenge[21].

Figure 8 presents a map showing the geographical distribution of the workshop locations and the participants' locations.

Figure 7: *Breakdown of workshop participants by organization Type*

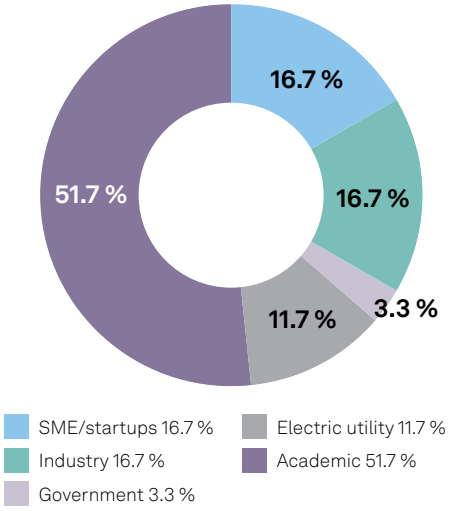
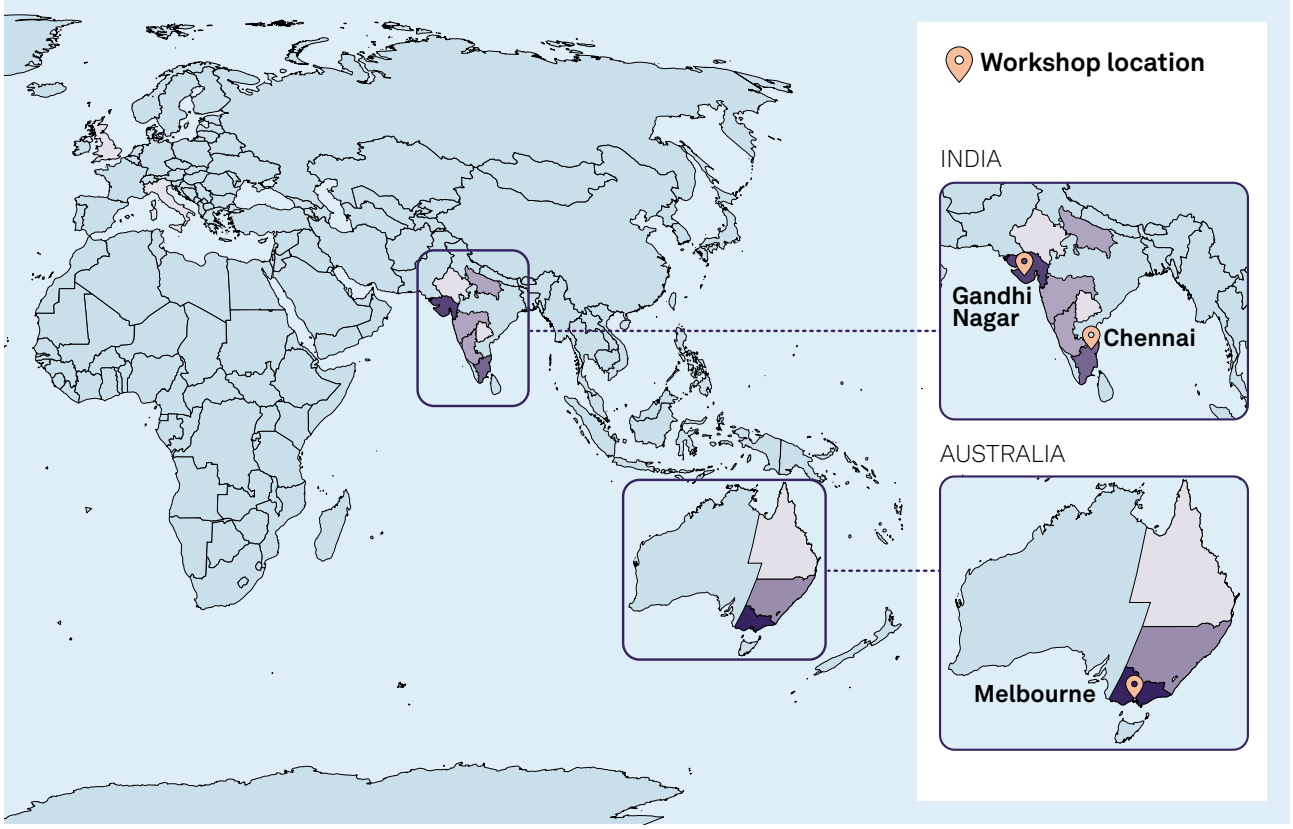


Figure 8: *Geographical distribution of workshop locations and participant counts*



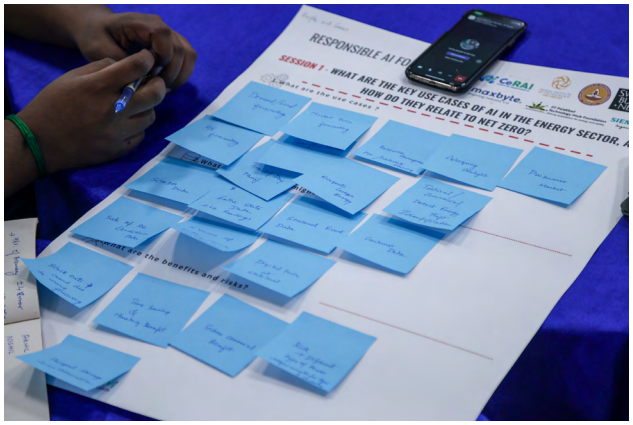
Workshop session themes

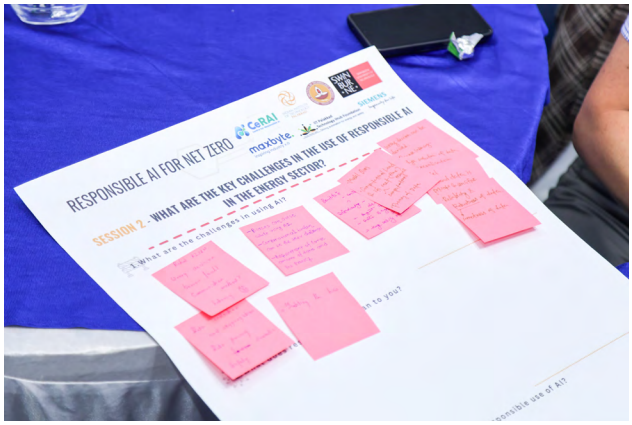
The sessions focused on the following key themes as described in Table 3. During the breakout sessions, participants worked in smaller groups to discuss and capture their ideas across each of the breakout session question area.

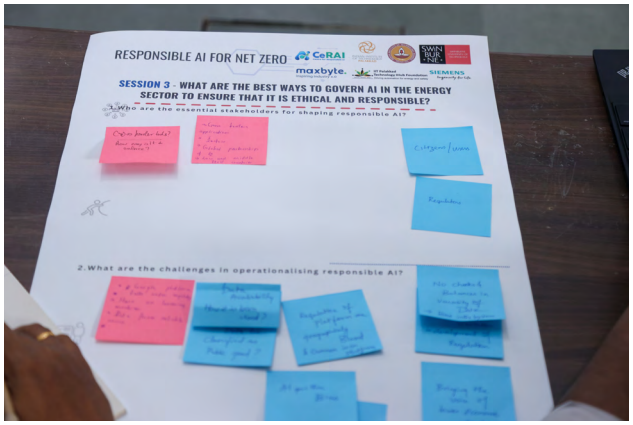
Table 3: Summary of objective of each session and questions used for breakout discussions

Session	Objective	Breakout questions
1	What are the key use cases of AI in the energy sector, and how do they relate to net zero?	– What are the use cases? – What is the data that might be used? – What are the benefits and risk?
2	What are the key challenges in the use of responsible AI in the energy sector?	– What are the challenges in using AI? – What does responsible AI mean to you? – What are the issues related to responsible use of AI?
3	What are the best ways to govern AI in the energy sector to ensure that it is ethical and responsible?	– Who are the essential stakeholders for shaping responsible AI? – What are the challenges in operationalising responsible AI?

Figure 10: *Photos of worksheets capturing key discussion points from each of the workshop sessions (Table 3)*







Foundations for establishing responsible AI for net zero guidelines

In this section, we present the outcomes from the analysis of data collected during the three workshops.

In Chennai, Gandhinagar and Melbourne the discussion was underpinned by two main assumptions:

1. First, when we are talking about AI we are talking about applications and use cases that are still emerging, with an interest in its potential for AI to help transform the energy sector.
2. Second, the transformation of the energy sector in support of sustainability was assumed to be vital and inevitable, with net zero simply one expression of sustainability.

These 'taken-for-granted' commitments to net zero meant that participants focused their discussions on *how* AI could be used to support that goal, instead of whether that was the right goal and technology (AI) for the energy sector.

The role of AI in supporting the transition to net zero

The first session of the workshops sought to uncover what the stakeholders felt was the role of AI in transitioning the energy sector towards net zero. By asking participants about use cases related to AI and energy, our aim was to explore what people knew about the use of AI in the energy sector, and how it could be used to support net zero.

In all three workshops participants demonstrated that they were aware of the possibilities and potential uses of AI and discussed the different ways that it could be used in the energy sector, including to support the sector's transition to net zero. This discussion drew on their existing knowledge of AI and of the energy sector. Based on this knowledge, there was a strong sense that AI could support efficiency and optimisation. In the workshops participants explored how these capabilities might be used across different parts of the energy sector value chain.

Key findings

AI is an enabler, not a solution

This perspective indicates that while AI can drive technological advancements, it is not a standalone solution but a tool that enables broader technological developments.

> *"We believe that AI is the enabler of new technologies. It's not the solution."*

Diverse applications and problem-specific AI

Participants acknowledged the complexity of selecting the appropriate AI system based on the specific problem to be addressed. This highlights the importance of problem definition in AI adoption.

> *"There are lots of different types of AI uniquely tailored to solving lots of different types of problems, and knowing the specific kind of problem that you have, is in itself a difficult thing to do, so that's something that needs to be considered going forward – what are we actually trying to solve here, because that will actually dictate the type of AI system that you actually want to implement."*

> *"Am I talking about controlling energy management at the transmission level? Am I talking about energy management at the solution level? Or, am I talking about energy management at the customer level?"*

Awareness of AI's potential

Participants of all three workshops were aware of the use of AI for prediction and accurate demand-based forecasting which could shift the energy sector from reactive to anticipatory management. Key examples include:

1. Predictive Maintenance: AI can anticipate maintenance needs for the grid, reducing infrastructure failures and operational disruptions.
2. Demand Management: AI can forecast energy demand and manage load, particularly for distributed energy resources like home batteries and rooftop solar panels.

> *"I think for predicting energy demand it's quite useful, because the data is already there."*

> *"You could use AI at an individual home level to shut off my refrigerator when I don't need it, shut off my AC when I don't need it, based on my consumption, and my society's consumption."*

> *"AI will be much more impactful to derive decisions where we have much more complex systems, where it is difficult to get some insights using the conventional approaches like mathematical optimisation."*

Responsible AI: risks, challenges and opportunities

Although not all participants had a detailed understanding or could define responsible AI, they consistently associated AI with risks, accountability, data, trust, security, and privacy throughout the workshops, often comparing AI to other technologies.

They understood that AI systems are built on and work because of the availability of digital data. For example, they drew on their understanding of data and algorithmic decision making to comment that a lack of accurate, locally relevant data would make it hard to train AI to anticipate what was likely to happen across specific Indian and Australian contexts. Security, privacy, and bias were all noted as particularly significant, with participants frequently posing the question 'what happens when AI makes incorrect predictions?'

Participants related *readiness* (for responsible use of AI) to the availability and accuracy of digital data, and to the people and processes that would be held accountable. Participants felt that neither data nor accountability structures were ready to support AI at present, commenting on the piecemeal nature of data and fragmented accountability in the energy sector in both Australia and India.

They noted that accountability is often not always clear, or easy to assign. This was related to the way that AI makes decisions, often within 'black box' systems, and complicated governance structures in both Australia and India. For example, in both Australia and India, governance of the energy sector is currently fragmented with different entities responsible for electricity generation, transmission, and sale to end-users. This is compounded by the roles of different State and Federal government agencies in balancing the needs of the sector with those of end-users.

Key findings

Data

Participants agreed that the impact of AI depends on the availability and quality of data. They noted that inadequate, incomplete, or inaccurate data poses significant challenges in selecting and implementing the right AI tools. Participants also highlighted that reliable data is vital for successful AI implementation. The challenge of sourcing data from the energy sector was related to:

- A reluctance within the sector to share commercially sensitive data
- The complexity of the sector and the different stakeholders involved which makes it hard to know the sector in its entirety using data
- Concerns about personal data and privacy, which extended to data sourced from the home, for example, through smart meters.

> *"Then comes the questions of do I have data? Where is that data? Is it complete? Is it right? Is it accurate? Then once you've gauged that, then we can comment on which AI tool you want to go with. Say, should it be supervised or unsupervised? Neural network? Deep neural network? The challenge starts with the availability of data, and then the reliability of the data."*

Country and context-specific data

An important issue identified was the lack of locally relevant data. Participants observed that AI models often rely on data from different contexts, such as Australian or US grid data, which may not accurately reflect energy usage patterns in India. While this was flagged as an operational issue, it became particularly acute as an issue that limited the ability to train AI to respond to local conditions.

> *"It has been my experience that when people talk about data in the energy sector they are all either Australian grid data sets or they are US grid data sets. There's no Indian grid data sets."*

> *"I think that one thing... is the lack of quality data at scale."*

Privacy, security and bias

Participants were also concerned about data privacy, bias, and security. They recognised that one source of data that would be important for the training and operation of responsible AI was personal data, including data drawn directly from technologies already located in the private space of the home e.g. smart meters, as well as metadata about people's behaviours. Access to this data is critical for tempering the bias that is built into AI systems that are trained using international data. But steps must also be taken to ensure that the personal data that is used is representative, so as to avoid biasing AI systems in favour of some end-users over others. It is also critical that access to those data be balanced with the right to privacy.

Participants also recognised that these biases could operate at a national level – for example when the data used to train AI was drawn from another country. At this level participants noted that bias represented a risk not only to privacy, but to the security of the grid itself. To address these concerns, they emphasised the need for robust cyber-security measures to be put in place to protect AI systems, including systems that were capable of detecting bias in data, and in the decisions made by AI using those data.

"Are these algorithms designed by toolchains in a way which benefits certain people? How do we ensure that they are fair? Should that not come under the regulation as well, because I think 90% of the AI today is open source."

"Do you know exactly what data has been taken? What sort of improvement this AI solution is achieving? And what is the possible funding of this AI solution? As far as I'm aware we don't have hardline standards across all institutions, and I think that is also a need."

"Bias has multiple levels right. Some are good bias, some are bad bias. We need a detection system that can detect and at least tell us."

Complexity of energy infrastructure

Participants acknowledged the risks in implementation of AI are not insignificant. However, for AI to be responsible it needs to actively contribute to positive outcomes. For this to happen participants recognised that we need to create an environment that enables AI to act responsibly. Participants emphasised that to create this environment:

- AI needs to work with the other systems, processes, stakeholders and technologies that already make up the energy sector
- That for this to happen those systems, processes, stakeholders and technologies would need to be ready to support AI.

They emphasised that if the energy sector focused on getting ready for AI it could address this fragmentation by considering where, when, and how AI should be used as part of the energy sector. This would allow the sector to move away from managing the adverse outcomes of AI systems, to proactively engaging AI and designing new AI systems that promote desirable outcomes.

"One of the other key challenges that we see is the complexity around integrating AI use cases into existing energy infrastructure."

"Generation to consumption is one train. You cannot say I will control here but I won't control here."

Computing power and carbon footprint

Participants raised concerns about the substantial computing power required by AI systems and its associated carbon footprint. They stressed the need to balance the efficiency gains of AI with its environmental impact, ensuring that AI contributes positively towards net zero goals without offsetting its benefits with high carbon emissions. This is a reminder to make sure that when we implement AI we do not simply pass the carbon cost down the supply chain. This insight is critical because it implies that for AI to support the goal of net zero responsibly, it must not only contribute to that objective but also actively avoid any adverse effects.

"When you have something that is deployed in the cloud how do you quantify the implications of implementing a new technology there? Because at the moment when you deploy something with AWS you don't get the carbon print so that's not taken into account."

"Are these AI models leading to net zero? Because it also consumes too much energy, which is a broad societal issue to discuss – how energy intensive these tasks are."

"Are you looking at estimating the impact of emissions on solar power generation."

Accountability and explainability

Accountability is a human process. It involves people who are accountable when AI makes mistakes, and people who are accountable for decisions about what AI should do. To ensure that those decisions can be consistently enforced, it is important that they are formalised through regulation. In this way, accountability creates an enabling framework within which AI can act responsibly. The opaque nature of AI decision-making and the fragmented governance of the energy sector complicates accountability. Participants felt that clear accountability is significant to establishing responsible AI use as well as AI's ability to explain the decisions.

Participants also noted that the question of what AI should do and how it should do that will change as the context in which AI is used changes. To account for what AI should do it is important to include diverse stakeholder voices in decisions about what AI should do. This should happen early on, when AI and its applications are being developed, and it should be ongoing, as part of regulatory processes.

"As with any new technology there will always be things that go wrong that we could never possibly foresee, and we need to be able to, when that happens, mobilise quickly and actually address those issues."

"With AI systems they can be a black box, so it can be hard to understand the logic and reasoning that took place... should AI be used to make decisions, or should it just provide advice? Where do you draw the line on this?"

Best practices and considerations for governing AI

Data readiness

As we get ready for the widespread adoption of AI in the energy sector it is imperative that we start to work towards a data landscape that:

- Helps the sector build and train AI that can act responsibly
- Can support AI making accurate and responsible decisions
- Can support the development of well curated, accurate and balanced (without biases) datasets that does not violate any privacy concerns

Participants highlighted that a robust data infrastructure to support AI implementation in the energy sector is essential. This also involves creating a data environment that enables AI to function responsibly and make precise decisions, thereby AI can contribute positively to the sector's goals.

> *"The success of AI lies in de-facto, correct, complete, accurate data."*

Accountability

Accountability is a human process. It involves people who are accountable when AI makes mistakes, and people who are accountable for decisions about what AI should do. To ensure that those decisions can be consistently enforced, it is important that they are formalised through regulation. In this way, accountability creates an enabling framework within which AI can act responsibly.

> *"When we talk of automation, automation needs to be precise, and perfect, and accurate. To be able to make that automation precise we want to build some intelligence. It can't be self-driven."*

Regardless of how well designed it is, participants recognised that like all technologies AI would make mistakes. While it was important to reduce the likelihood of mistakes being made, participants expressed the importance of clarity about who was accountable when it occurred. In addition to accountability for technical failures, this also includes accountability for instances where AI is used in ways that do not comply with regulations. The implication here is that AI should not only be used transparently but also be understandable in terms of how it functions.

Accountability is essential for responsible implementation of AI in the energy sector. It involves identifying the right people who will be held accountable when AI makes mistakes and for decisions about what AI should do. While AI can be involved in decision making processes, workshop participants recognised that people are ultimately accountable for what AI does and how it is used in the energy sector. Formalising accountability through regulation is necessary to lead to consistent enforcement and responsible AI behaviour. That accountability happens at three different points in AI's lifecycle:

- When the technical capacities of AI are decided during its design and development
- When decisions are made about if and how AI should be used in the energy sector
- When AI makes a mistake

Participants noted that while AI can be used to make decisions about the day-to-day operation and maintenance and operation of the energy sector, the decisions about what AI does still need to be made by people. This was particularly important in the context of net zero. Participants emphasised that people were accountable for this goal, and for keeping AI on track by building AI that works towards that goal, and through the ongoing and effective management of that technology.

> *"Another aspect which I feel is that even with the responsible use of AI there is a scenario where a fault has occurred. Who is to blame?"*

> *"Humans can make decisions and be held accountable, but with an AI system you have the dilemma of who should be held accountable for the repercussions of the systems."*

Regulation

Regulation formalises accountability and helps to enforce accepted standards to avoid bias or breaches of privacy, for example, along with other established ethical AI principles. Effective regulation can guide consistent and contextually appropriate decision-making and ensuring safe application relevant to the energy sector. However, participants noted that the use of AI in the energy sector was currently unregulated, which meant that formal structures to support accountability are not yet in place. This formalisation is vital, because it reduces risks, builds in consistent standardised processes, and establishes contextually appropriate decision making that considers the needs of the diverse stakeholders most likely to be impacted by the use of AI.

> *"Our number one issue was the bias... so biases can arise in the models, which are made by humans, and also the data which is fed... and that can also lead to biased outcomes."*

> *"And we found that critical infrastructure regulation in Australia today doesn't really take into effect what the impact of AI will be on regulated infrastructure."*

Participants did acknowledge that regulation could be difficult to get right because it involved balancing rules that were applied consistently across different jurisdictions, with rules that were designed for local conditions. The need to respond to local conditions was particularly important, because people recognised that working responsibly in support of net zero would look different in different locations based on environmental conditions, the design, operation, and governance of the energy sector, and the geopolitical context.

For example, in India AI would need to operate across different states with energy sectors that were structured and funded differently. Contrasting this, while there was still a need to navigate a complex relational landscape to regulate responsible AI, the relatively limited geopolitical complexity in Australia meant there was a sense that that was achievable. As a result, the question for participants was not if Australia could implement responsible AI in support of net zero, but how it could become a leader in that space.

Guidelines for translating ethical principles into practical responsible AI energy governance

AI technologies can revolutionise smart energy systems, which is pivotal for achieving net zero goals and fostering sustainable energy practices.

However, there is less clarity on the essential aspects required to effectively implement responsible AI principles. This gap hinders stakeholders from fully embracing these principles. The complexity of integrating AI technologies also presents a significant barrier.

However, without a clear roadmap for implementing responsible AI principles, stakeholders may struggle to identify and prioritise the key aspects necessary for successful AI adoption in current energy systems. A well-defined implementation strategy is essential.

Based on the analysis of the inputs received from various stakeholders across the three workshops, and analysis gathered from the literature, the key guidelines for translating ethical principles into responsible AI energy governance practice can be summarised as follows:

Democratise data collection and development of AI models to enable unbiased and explainable AI systems

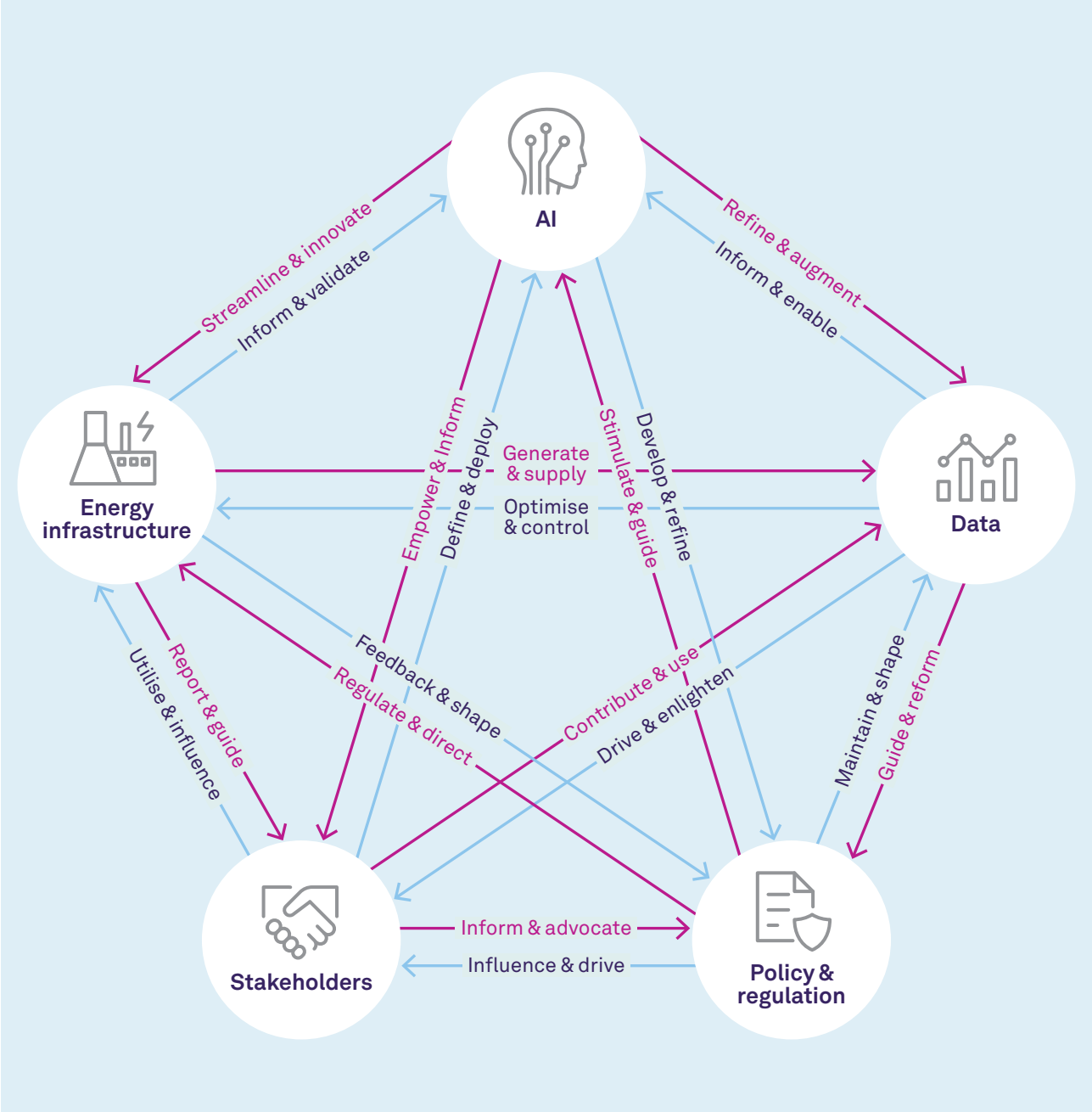
- Build accountability of AI systems by making people accountable for how AI outcomes are applied to energy systems, and to handle situations when AI makes mistakes
- Introduce regulations that formalise accountability around the use of AI in the energy sector

To support the transition to implementing sector-wide responsible AI, we propose as a roadmap (see Figure 11) that outlines five pillars and their collaborative practices that guide the implementation of the responsible AI principles within the energy sector.

The roadmap is designed to guide the implementation of responsible AI principles in energy systems through collaborative practices. In conclusion, we emphasise that operationalising responsible AI in the energy sector requires ongoing efforts and strategic actions. The path towards responsible AI-driven smart energy systems demands collaborative efforts and a collective commitment to ethical AI innovation, ensuring that technological advancements contribute to a sustainable and equitable energy future.

Pillar	Description
AI 	AI represents the AI technologies that interpret, analyse and predict energy-related data to optimise the performance of energy systems.
Data 	Data represents the information foundations of energy systems. It focuses on the acquisition, management and consumption of data related to energy generation, distribution, transmission, and consumption.
Stakeholders 	Stakeholders form a collaborative hub, acknowledging the multitude of participants within the energy sector. This pillar recognises the diverse set of actors involved in the energy system, from producers to consumers, and their roles in transition to sustainable practices. These stakeholders include energy companies, utility providers, technology developers, policymakers, and end-users, all of whom have a stake in the development and implementation of AI in the energy sector.
Policy and regulation 	Policy and Regulation represent the critical role of a governance framework that ensures AI applications are developed and utilised within ethical and regulatory boundaries. Such a framework is not merely prescriptive but evolves through continuous dialogue with technological advancements and stakeholder feedback.
Energy infrastructure 	Energy Infrastructure serves as the operational backbone of energy systems, which include digital platforms capable of supporting the sophisticated requirements of AI-driven smart energy systems.

Figure 11: *Roadmap for establishing responsible AI-driven smart energy system to achieve net zero goals*



Summary

As nations grapple with climate change and look to create efficiency and sustainability in the energy sector, AI solutions have gained prominence as tools that stakeholders in both Australia and India agree have the potential to impact net zero targets.

We addressed the importance of how AI is implemented in the energy sector, emphasising that for AI to have a meaningful impact on net zero targets, and to ensure that AI is responsible, it needs to be able to work as part of the configuration of the other technologies, systems, and processes that already make up the energy sector. We showed that it is when we account for these relationships, as well as the importance of people and regulation in ensuring that it is accountable, that AI becomes responsible.

Like the energy sector to which it will be introduced, the relationship between AI, responsibility, and net zero is complex. As participants pointed out in all three workshops, this should not prevent us from using AI to achieve net zero targets in that sector. Instead, it is something we must be sensitive to when we design and deploy AI, being careful to establish a relationship between AI and the energy sector that is reciprocal, with each informed by the other.

Our findings are informed by our analysis of the three interactive workshops with energy sector stakeholders and subject matter experts (in Chennai, Gandhinagar and Melbourne) and systematic desktop literature review. Key insights from our engagements with stakeholders in the energy sector, which shape our understanding of this relationship, include:

AI's promise and potential

- People in the energy sector are aware of and able to easily discuss the use of AI in the energy sector, and in support of net zero.
- People understand AI primarily as a tool that can be used to support prediction and forecasting, for example predictive maintenance of grid infrastructure.
- AI technologies can revolutionise smart energy systems, which is pivotal for achieving net zero goals and fostering sustainable energy practices.

AI introduces risks

- Using AI entails risk, but the responsible use of AI is about more than simply reducing that risk. It is about actively contributing to positive outcomes.
- For AI to work effectively the systems, process, and technologies that make up the energy sector need to be ready to support AI. The complexity of integrating AI technologies into a complex energy infrastructure also presents significant risks.
- Data is particularly important to ensure that AI makes decisions that are accurate, privacy-preserved and unbiased.
- AI while introducing several risks, there is growing interest and acceptance in the potential use of AI to help transform the energy sector.

Summary

Essential responsible AI principles

- Like any technology, AI will make mistakes. Accountability for those mistakes (and for the decisions that AI makes more broadly) needs to be clearly attributable to people.
- AI needs to be able to respond to local conditions, while also working consistently across different contexts.
- Decisions about how AI is used should be developed in dialogue with diverse stakeholders from within and outside the energy sector.
- Understanding of the essential aspects and principles required to effectively implement and operationalise responsible AI principles is key to bridge the gap that hinders stakeholders from fully embracing AI.
- Seven key dimensions of responsible AI namely – Data, Algorithm, Human-Centric, Social and Societal, Governance and Regulatory, Environmental, and Economic and Labour are crucial for promoting ethical, equitable, and sustainable AI deployment in the energy sector.

Need for clear roadmap

- Data and a detailed analysis of needs and goals was identified as vital for supporting the development of regulation that responded to local conditions
- Without a clear roadmap for implementing responsible AI principles, stakeholders may struggle to identify and prioritise the key aspects necessary for successful AI adoption in current energy systems. A well-defined implementation strategy is essential.
- The roadmap we present is a starting point that can contribute to the development and integration of responsible AI principles and practices into the energy sector aimed towards a more resilient, sustainable and environmentally friendly energy future supporting net zero.
- Further work is needed, involving multi-stakeholder collaboration involving government, industry, academia, energy regulators and startups to realise responsible AI across the sector.

References

1 Marcello B. Solomon, Swapna S. Rabha et al. Decarbonization in Australia and India: Bilateral Opportunities and Challenges for the Net Zero Transformation, ACS Engineering Au 2024 4 (3), 295-311 DOI: 10.1021/acsengineeringau.3c00060

2 Qamar Akhter, Abubakar Siddique, Salman A Alqathani, Anzar Mahmood, Mehboob Alam, Zohaib Mushtaq, Muhammad Farrukh Qureshi, Waseem Aslam, and Paranav Kumar Pathak. 2023. Efficient energy management for household: Optimization-based integration of distributed energy resources in smart grid. IEEE Access (2023)

3 Jatin Bedi and Durga Toshniwal. 2019. Deep learning framework to forecast electricity demand. Applied energy 238 (2019), 1312–1326

4 Latifa A Yousef, Hibba Yousef, and Lisandra Rocha-Meneses. 2023. Artificial Intelligence for Management of Variable Renewable Energy Systems: A Review of Current Status and Future Directions. Energies 16, 24 (2023), 8057

5 Amirhossein Ahmadi, Mojtaba Nabipour, Behnam Mohammadi-Ivatloo, Ali Moradi Amani, Seungmin Rho, and Md Jalil Piran. 2020. Long-term wind power forecasting using tree-based learning algorithms. IEEE Access 8 (2020), 151511–151522

6 Jinendra Rahul, Yagvalkya Sharma, and Dinesh Birla. 2012. A new attempt to optimize optimal power flow based transmission losses using genetic algorithm. In 2012 Fourth International Conference on Computational Intelligence and Communication Networks. 566–570

7 CR Balamurugan. 2018. Three area power system load frequency control using fuzzy logic controller. International Journal of Applied Power Engineering (IJAPE) 7, 1 (2018), 18–26

8 Shahriar Rahman Fahim, Subrata K Sarker, SM Muyeen, Sajal K Das, and Innocent Kamwa. 2021. A deep learning based intelligent approach in detection and classification of transmission line faults. International Journal of Electrical Power & Energy Systems 133 (2021), 107102

9 Dibya Bharti. 2020. Multi-point Optimal Placement of Shunt Capacitor in Radial Distribution Network: A Comparison. In 2020 International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET). 1–6

10 Vasudevarao Kothakotla and Bharath Kumar. 2021. Analysis and Design of Robust PID Controller for Grid Voltage Control of Islanded Microgrid Using Genetic Algorithm. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). 719–726

11 Yixing Wang, Meiqin Liu, and Zhejing Bao. 2016. Deep learning neural network for power system fault diagnosis. In 2016 35th Chinese control conference (CCC). 6678–6683

12 Iker Esnaola-Gonzalez, Marko Jelić, Dea Pujić, Francisco Javier Diez, and Nikola Tomašević. 2021. An AI-powered system for residential demand response. 10, 6 (2021), 693

13 Yantai Huang, Hongjun Tian, and Lei Wang. 2015. Demand response for home energy management system. International Journal of Electrical Power & Energy Systems 73 (2015), 448–455

14 Haosen Qin, Zhen Yu, Tailu Li, Xueliang Liu, and Li Li. 2023. Energy-efficient heating control for nearly zero energy residential buildings with deep reinforcement learning. Energy 264 (2023), 126209

15 Yonghua Zhu, Xinqiao Jin, and Zhimin Du. 2012. Fault diagnosis for sensors in air handling unit based on neural network pre-processed by wavelet and fractal. Energy and buildings 44 (2012), 7–16

16 Asif Iqbal Kawoosa, Deepak Prashar, Muhammad Faheem, Nishant Jha, and Arfat Ahmad Khan. 2023. Using machine learning ensemble method for detection of energy theft in smart meters. IET Generation, Transmission & Distribution 17, 21 (2023), 4794–4809

17 Roba Alsaigh, Rashid Mehmood, and Iyad Katib. 2023. AI explainability and governance in smart energy systems: A review. Frontiers in Energy Research 11 (2023). https://doi.org/10.3389/fenrg.2023.1071291

18 Ioannis Antonopoulos, Valentin Robu, Benoit Couraud, Desen Kirli, Sonam Norbu, Aristides Kiprakis, David Flynn, Sergio Elizondo-Gonzalez, and Steve Wattam. 2020. Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. Renewable and Sustainable Energy Reviews 130 (2020), 109899. https://doi.org/10.1016/j.rser.2020.109899

References

19 Tanveer Ahmad, Dongdong Zhang, Chao Huang, Hongcai Zhang, Ningyi Dai, Yonghua Song, and Huanxin Chen. 2021. Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities. Journal of Cleaner Production 289 (2021), 125834. https://doi.org/10.1016/j.jclepro.2021.125834

20 Jillian Carmody, Samir Shringarpure, and Gerhard Van de Venter. 2021. AI and privacy concerns: a smart meter case study. Journal of Information, Communication and Ethics in Society 19, 4 (2021), 492–505. https://doi.org/10.1108/JICES04-2021-0042

21 E. Bomasang, in Advancing Women's Representation and Leadership in the Energy Sector vol. 2023, ed: https://www.abtglobal.com/, 2023

Glossary of terms

NASSCOM: National Association of Software and Service Companies.

Net zero: Net zero refers to the balance between the amount of greenhouse gas (GHG) that is produced and the amount that is removed from the atmosphere.

Greenhouse gag (GHG): Greenhouse gases are gases which trap heat in the earth's atmosphere and contribute to climate change.

Artificial Intelligence (AI): Artificial intelligence is the simulation of human intelligence processes by machines or computer systems.

Smart energy systems (SESs): SESs integrate various energy sectors, including power (electricity), heating, cooling, and transportation, in an intelligent manner.

Responsible AI: Responsible AI operationalises ethical principles into tangible strategies, guiding AI deployment in real-world scenarios.

Institute of Electrical and Electronics Engineers (IEEE): The Institute of Electrical and Electronics Engineers is an American professional association for electronics engineering, electrical engineering, and other related disciplines.

Prosumer: Energy consumers, often called prosumers because they both consume and produce electricity.

Distributed energy resources (DERs): DERs are small-scale units of local power generation, typically located close to where the electricity is used (e.g., homes, businesses). These resources can include solar panels, wind turbines, battery storage systems, and other forms of renewable or conventional energy.

Load forecasting: Load forecasting predicts future electricity demand. It can be short-term, medium-term, or long-term, depending on the time frame of the prediction.

Generation forecasting: Generation forecasting predicts the amount of electricity that power plants will produce over a specific period.

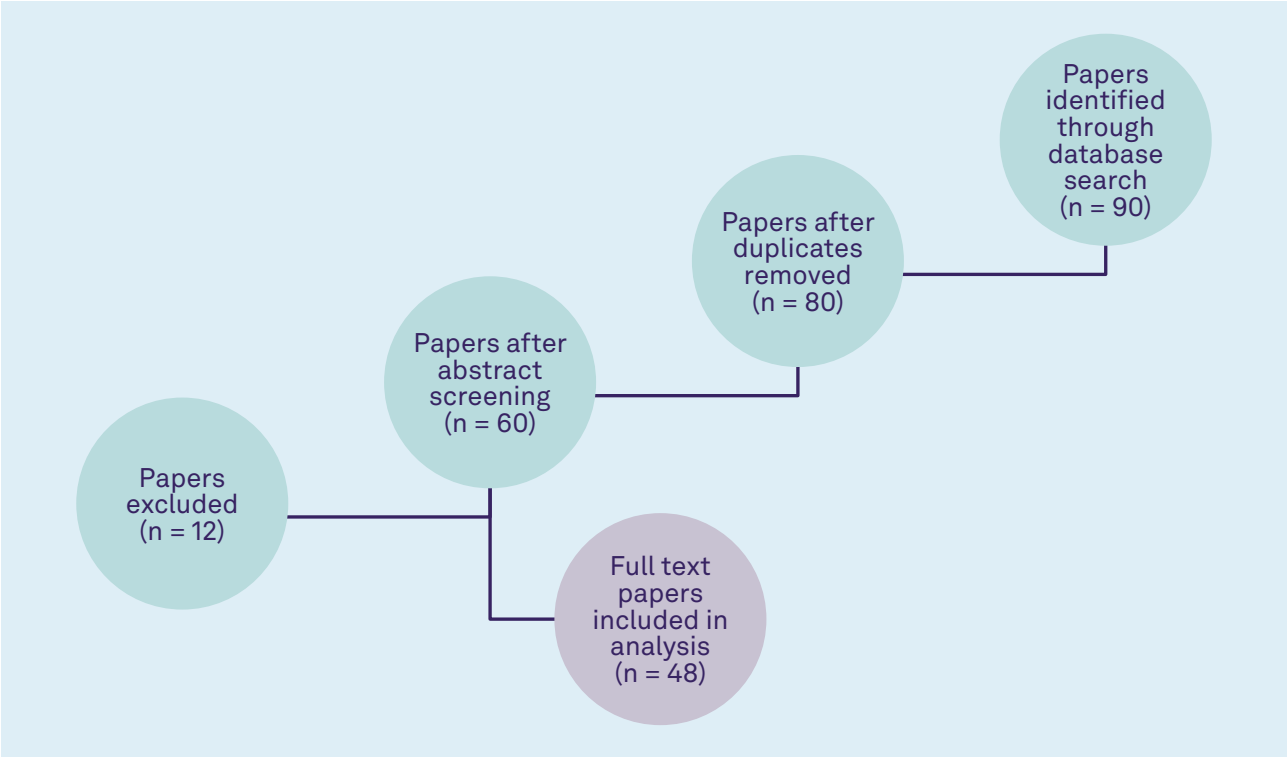
Demand response management: demand response management involves adjusting electricity consumption by end-users during peak demand periods or when energy supply is limited.

Small medium-sized enterprises (SMEs): Small and medium-sized enterprises (SMEs) are businesses with a limited number of employees and relatively low revenue compared to larger companies.

Predictive maintenance: Predictive maintenance involves monitoring the condition of equipment and systems to predict when maintenance should be performed.

Appendix A – Systematic literature review

Figure 12: *Roadmap for establishing responsible AI-driven smart energy system to achieve net zero goals*



Research question

- How does the use of AI impact the energy sector?
- What are the current ethical principles relevant to the adoption of AI in the energy sector?
- What are the key ethical issues surrounding the use of AI in developing smart energy systems to achieve 'net zero' targets?
- How can ethical principles be effectively translated into AI energy governance practice?

Search keywords

- AI in energy
- Net Zero AND AI ethics (responsible AI)
- AI ethics in energy
- AI ethics in the energy sector
- AI ethics for sustainable energy
- AI ethics in smart energy systems
- Energy equity and affordability and AI
- Responsible (ethical) AI for energy infrastructure
- Ethical AI energy framework
- Ethical smart energy systems
- Ethical AI governance in the energy sector

Information sources

- Google scholar
- Web of Science
- IEEE Xplore
- ACM Digital Library
- ScienceDirect
- arXiv

Eligibility criteria

- Inclusion
  - Papers published between 2020 and 2023
- Exclusion
  - Paper that did not meet our research question based on abstract screening and subsequently full paper review
  - Articles from websites without reasonable evidence of peer review or editorial review

Appendix B – Research Workshops in Australia and India

Research workshop 1

Venue: IIT Madras, India
Date: 8th Dec 2023





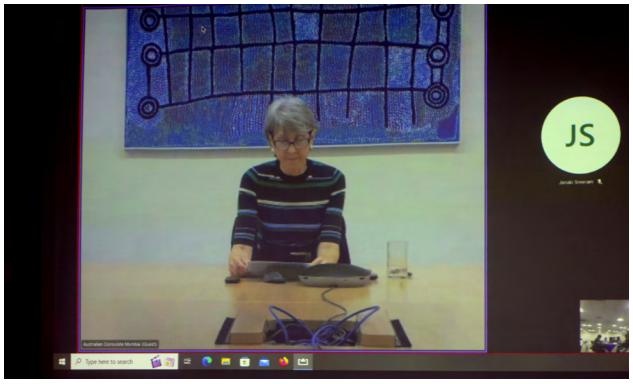




Appendix B – Research Workshops in Australia and India

Research workshop 2

Venue: IIT Gandhinagar, India
Date: 15th Dec 2023















Appendix B – Research Workshops in Australia and India

Research Workshop 3

Venue: Radisson Melbourne, Australia
Date: 12th Mar 2024







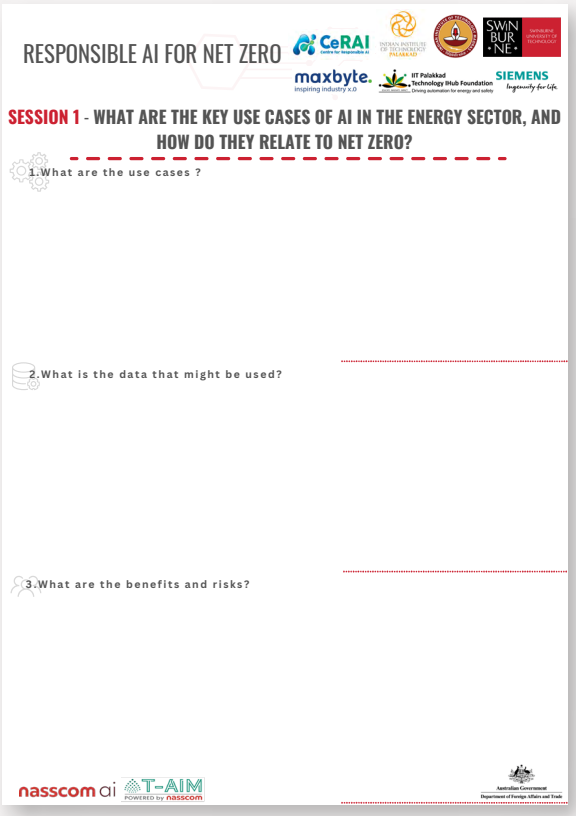








Appendix C – Research workshop worksheets

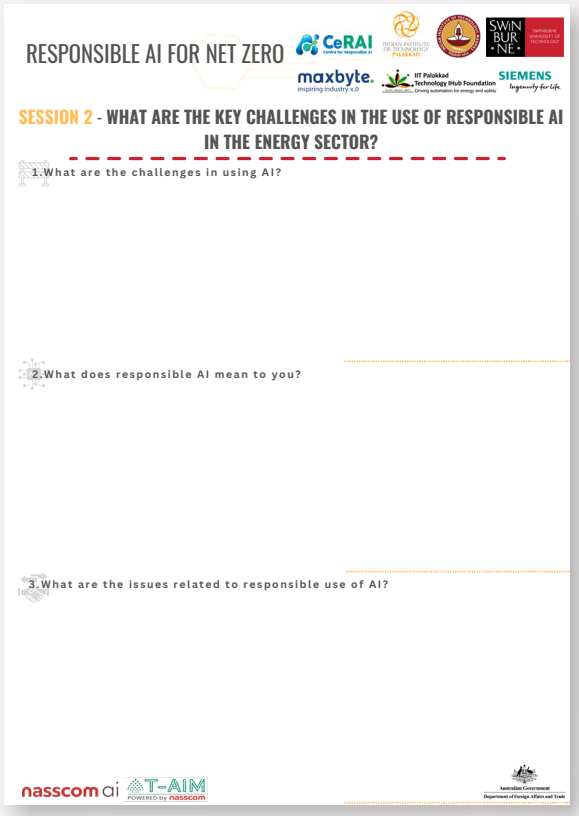

RESPONSIBLE AI FOR NET ZERO

SESSION 1 - WHAT ARE THE KEY USE CASES OF AI IN THE ENERGY SECTOR, AND HOW DO THEY RELATE TO NET ZERO?

1. What are the use cases ?

2. What is the data that might be used?

3. What are the benefits and risks?

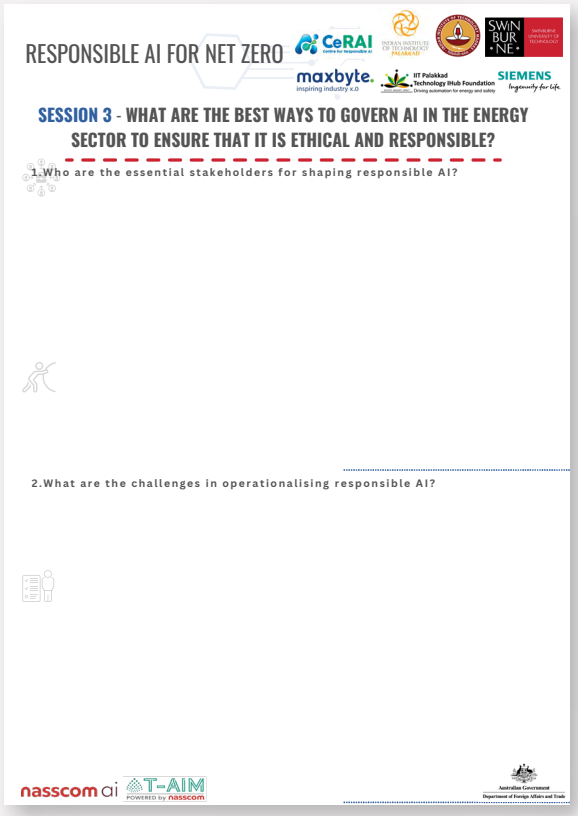

RESPONSIBLE AI FOR NET ZERO

SESSION 2 - WHAT ARE THE KEY CHALLENGES IN THE USE OF RESPONSIBLE AI IN THE ENERGY SECTOR?

1. What are the challenges in using AI?

2. What does responsible AI mean to you?

3. What are the issues related to responsible use of AI?


RESPONSIBLE AI FOR NET ZERO

SESSION 3 - WHAT ARE THE BEST WAYS TO GOVERN AI IN THE ENERGY SECTOR TO ENSURE THAT IT IS ETHICAL AND RESPONSIBLE?

1. Who are the essential stakeholders for shaping responsible AI?

2. What are the challenges in operationalising responsible AI?


SWIN
BUR
NE
SWINBURNE
UNIVERSITY OF
TECHNOLOGY