TRILATERAL RESEARCH
Ethical AI

RAiUK

University of Nottingham
UK | CHINA | MALAYSIA

UK Research and Innovation

RAISE

Responsible GenerativeAI for SMEs in UK and Africa

RAISE Guidelines

Version 1.0

CONTENTS

User Guide

This booklet contains the first version of our guidance on the responsible use of generative AI for SMEs in the UK and Africa. This document was developed by members of the RAISE project, an impact accelerator project funded by RAI UK.

The purpose of this document is to collect feedback on content and representation of these guidelines. We would therefore greatly appreciate it, if you could tell us what you like or dislike about these guidelines. Do you feel they would be useful to you? What would they need to become more relevant and applicable? Do they contain content that you would not need? We welcome any sort of feedback.

You can provide feedback and download a pdf version of this document at:

https://forms.office.com/e/RPS6Xs3hLi

SCAN TO PROVIDE FEEDBACK

1. INTRODUCTION

In the last decade, generative Artificial Intelligence (AI) has kick-started a new wave of AI integration and adoption. These range from systems for text generation, translation, and coding (such as OpenAI's ChatGPT based on Generative Pre-trained Transformer models, Google's Gemini), for image generation (e.g NVIDIA's StyleGAN (Generative Adversarial Network) and Google's DeepDream) and video generation (e.g OpenAI's Sora). The power of these systems has led to a flurry of activities to develop and support them, and they offer significant commercial opportunities in the form of new business models and exploitation. Generative AI systems have the potential to contribute to human flourishing by enhancing creativity and innovation, efficiency, and productivity, facilitating economic growth and sustainability, providing inclusive solutions, and ensuring social responsiveness. At the same time, they raise numerous concerns, including their impact on the educational system, accuracy, intellectual property (e.g., copyright) infringement, data governance, liability, and transparency.

Small to Medium Enterprises (SMEs), including technology start-ups, are driving the development of Generative Artificial Intelligence products and services, whilst some non-technology SMEs are enthusiastically adopting these tools. However, SMEs do not typically have the in-house resources or specialists for responsible AI deployment, particularly in

fast-moving areas with uncertain regulation and guidance. Further, SMEs are embedded in commercial ecosystems and therefore can be at the mercy of larger providers. In the case of generative AI, SMEs will often be deploying tools on other parties' terms and conditions.

Beyond the described challenges experienced within the UK generative AI context, SMEs in Africa operate within diverse languages, cultures, and traditions. How generative AI systems align with or are sensitive to socio-cultural needs, contexts and expectations is critically important to the global discourse on ethical AI.

Explicitly, these guidelines consider:

1. the lack of control by SMEs over the generative AI tools used, including lack of access to data collection and validation at scale.

2. the developing nature of AI ethics guidance and the lack of its specificity and/or tailoring to the SME research and software development context.

3. the extra effort and time required by SMEs to ensure responsible AI practices.

4. the impact of local culture, norms, and practices on AI use future challenges related to regulatory changes.

5. the pace of developments in generative AI

The following are the priority actions elaborated within this document:

- **Section 1** introduces the target audience of the guidelines and explores potential beneficiaries beyond SMEs, discussing recent developments prompting the need for further guidance. It also delves into specific issues concerning SMEs in the realm of generative AI, with insights into potential pitfalls. Additionally, it examines the implications within both the United Kingdom (UK) and African contexts.

- **Section 2** focuses on the primary challenges and opportunities encountered in the integration of Artificial Intelligence (AI) within SMEs. It provides an overview of the current AI landscape within SMEs, outlining the global and contextual hurdles they face, while also highlighting the transformative potential of adopting generative artificial intelligence (AI).

- **Section 3** delineates the risks associated with AI adoption by SMEs, offers strategies for mitigating these risks, and discusses future development prospects for SMEs in this domain.

- **Section 4** proposes potential policies and approaches for the AI lifecycle that SMEs can adopt to ensure the responsible utilisation and seamless integration of AI solutions.

1.1. Who are the guidelines for?

We have put together this guidance for small to medium enterprises (SMEs). In the UK this is defined as any organisation with fewer than 250 staff, an annual turnover under €50m, and a balance sheet total under €43m[1]. It could include everything from a small law firm to a restaurant, to a boutique design firm.

The SME we have in mind:

- Has heard about the potential of generative AI tools but is not an expert.
- Is interested in exploring generative AI tools or adopting them for business use.
- Wants to make sure that they use generative AI in a safe and responsible way that will be best for them, their reputation, their customers and partners, and for wider society.

1.2. Who else might benefit from them?

Many not-for-profit organisations and charities face some of the same concerns as small business (e.g., trying to be effective and do the right thing, but with limited resources) and as such may also find this guidance useful.

Some technology start-ups are experts on generative AI. They're actively developing new generative AI tools and services based around them. These companies might be new and have a small number of

1 https://www.gov.uk/government/publications/beis-small-and-medium-enterprises-sme-action-plan-2022-to-2025/beis-small-and-medium-enterprises-smes-action-plan-2022-to-2025-accessible-webpage#definition-of-an-sme

employees, but they know generative AI inside out. These companies are not the primary audience for this guidance, but we think they may get something from it too, especially if they have an interest in responsible and ethical use, or their customers include other SMEs.

Many larger organisations are currently trying to understand how to use generative AI and may be looking to do so in a responsible and ethical way. In contrast to SMEs, they are more likely to have existing data protection or compliance teams and dedicated legal and technical expertise. However, generative AI is a new area, and even large organisations may benefit from this guidance.

AI ethics and AI governance are professions that are starting to emerge. We hope this guidance would also be useful for these professionals where they work with or support SMEs.

1.3. What are the guidelines about?

1.3.1. What are recent developments calling for additional guidance?

'Generative' means able to create. Generative AIs are software tools that are able to create new content and media, including text, images, video and audio.

It can be thought of as having four core use cases[2]:

- **Creation.** A user will typically give the AI tool a 'prompt' such as 'create me a job advert for an apprentice chef, to work in a newly opening restaurant' and the tool will use this to generate such an advert.

- **Concision.** A user provides a prompt that links to a text or includes a text, and the Gen AI provides a summary, whether for analysis or as a basis for external-facing text.

- **Co-piloting.** A system interacts with the user or with the general public in order to guide activities, such as research or customer service routing.

- **Coding.** The AI provides computer code based on natural language prompts.

In Watsonville, California, a Chevrolet dealership integrated a ChatGPT powered chatbot. A customer, Chris Bakke in a post on X (formerly Twitter) revealed how he got the chatbot to agree to sell a 2024 Chevy Tahoe (original cost of $76,825) for $1 in a supposedly legally binding deal. Under this post, one other user shared how the chatbot (for a Chevrolet dear) recommended a Ford F-150. Although the dealership did not honour the deal struck by the chatbot, this shows the risks of non-robust integration of generative AI systems.

2 https://www.mckinsey.com/industries/real-estate/our-insights/generative-ai-can-change-real-estate-but-the-industry-must-change-to-reap-the-benefits

The systems can carry out these tasks because they have been trained on a huge number of documents or images and are able to predict what the expected content, format, style, language and tone should be for their outputs. They work best when the desired output is well represented in the data used to train the model.

Some of the big changes with generative AI, when compared to previous AI and data science approaches are:

- **Accessibility** – It is comparatively easy to get started with using generative AI, at least in some manner. Some of the bigger-value uses require specialist expertise, but some tools are simply available to the public via a website or app, or may become integrated with other software on phones, or through other software.

- **Ability to work with unstructured data** – formerly, AI tools needed well collected, carefully processed and formatted data (think of data in a spreadsheet). Generative AI can work with 'unstructured' data such as existing documents designed to be read by people.

- **Ability to ask questions and create content using natural human language** in manner that can be close to having a conversation with a person.

- Potentially allow for **automation of higher level technical and creative fields** (such as creating computer code, business writing or digital images).

1.3.2. What are specific issues for SMEs?

We're still finding out more about the specific issues for SMEs in adopting generative AI, but we can build on what is known about the technology, the general issues facing SMEs and from other areas of technology adoption and responsible technology use, such as data protection for example. SMEs may find:

- **Lack of specialists** – a large company might have an R&D department, dedicated data protection lawyers and an AI governance team. An SME is unlikely to have such roles, meaning that AI adoption by SMEs is likely to be driven by generalists, with lots of other things to think about.

- **Lack of resources or tight margins** – SMEs often need to make every penny count.

- Technology, data processing or AI may not be central to their business model. AI might therefore be a **very new area**.

- **Unpredictability** of how an emerging technology will affect each market sector. Analysts make over-arching predictions that generative AI will have profound effects "on the market". A challenge for SMEs is to adopt and respond to the technology in ways that properly embrace the opportunities that the technology creates, without taking on undue risk.

On the other hand, SMEs can be great innovators with the drive to try out new ways of doing things. This is an area where we really want to check our assumptions.

1.3.3. UK and Africa

Small and medium enterprises (SMEs) in both the United Kingdom and Africa are increasingly embracing artificial intelligence (AI) technologies, including generative AI, to drive innovation, enhance productivity, and gain a competitive edge in their respective markets. However, as these SMEs integrate generative AI into their operations, they face unique challenges and considerations that necessitate clear guidelines for responsible use.

A major reason SMEs in both regions require guidelines for responsible AI use is to navigate the ethical implications of generative AI. Generative AI has the capability to produce highly realistic content, such as images, text, and music, autonomously. While this technology presents numerous opportunities for SMEs, it also raises concerns about ethical dilemmas, including potential misuse, manipulation, and infringement of intellectual property rights. Clear guidelines can help SMEs manage the ethical implications of using generative AI and provide them with practical ways to ensure transparency, accountability, and fairness in their AI applications.

Also, guidelines for responsible AI use are essential for SMEs in the UK and Africa to address regulatory requirements and compliance challenges. Both regions have distinct legal frameworks and regulations governing the use of AI technologies, such as data protection laws and intellectual property rights regulations (IPR). Many

African countries are in the process of developing or updating their regulatory frameworks for AI technologies, including data protection laws, intellectual property rights regulations (IPR), and guidelines for AI ethics and governance. SMEs must navigate these complex regulatory landscapes to avoid legal pitfalls and ensure their AI deployments are compliant with local and international standards. By providing SMEs with clear guidance, this can help mitigate legal risks and foster a culture of responsible AI adoption.

Furthermore, guidelines for responsible AI use can empower SMEs to build trust with customers, partners, and stakeholders. By demonstrating a commitment to ethical AI practices, SMEs can enhance their reputation and credibility in the marketplace. This trust not only strengthens relationships with existing stakeholders but also opens doors to new business opportunities and partnerships, driving long-term success and sustainability for SMEs in both the UK and Africa.

Therefore, by providing clear guidance on ethical considerations, regulatory compliance, and building trust, these guidelines can empower SMEs to harness the transformative potential of generative AI while mitigating risks and contributing to the responsible development and deployment of AI technologies in their respective regions.

2. ARTIFICIAL INTELLIGENCE FOR SMES: KEY CHALLENGES AND OPPORTUNITIES

2.1. Overview

What is the current state of AI in SMEs and what are the potential global and regional challenges and opportunities around the adoption of generative AI by SMEs?

2.1.1. What is the current state of AI in SMEs?

Despite the significant interest and hype surrounding the utilisation of AI, its integration into businesses remains in its early stages. A survey involving 200,000 businesses, conducted by the US Census Bureau, reveals that only a small fraction (3.8%) of industries have implemented AI to develop their products and services[3]. This contrast underscores the disparity between the current discourse on AI implementation and its actual use in business operations[4]. It also indicates that SMEs

3 https://www.census.gov/library/stories/2023/11/businesses-use-ai.html

4 https://www.nbcnews.com/data-graphics/wide-gap-ais-hype-use-business-rcna127210

interested in integrating generative AI into their businesses exhibit a degree of reluctance to do so, and it is possible to speculate that this might particularly be due to the lack of clear guidelines that could serve as a reference point for how to do so responsibly.

2.1.2. What are the frameworks for responsible AI?

Currently, several global guidelines reflecting various principles for responsible AI adoption exist. These include:

- Responsible AI Adoption: 5 Steps Every Startup Must Take[5]
- OECD AI Principles[6]
- Trustworthy AI principles[7]
- EU AI Act[8]
- UK Algorithmic Transparency Recording Standard[9]
- Edinburgh declaration on responsibility for responsible AI[10]
- UNESCO Guidance for generative AI in education and research[11]

5 https://yourstory.com/2023/11/startup-guide-ethical-ai-practices

6 https://oecd.ai/en/ai-principles

7 https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

8 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

9 https://www.gov.uk/government/publications/algorithmic-transparency-template

10 https://medium.com/@svallor_10030/edinburgh-declaration-on-responsibility-for-responsible-ai-1a98ed2e328b

11 https://unesdoc.unesco.org/ark:/48223/pf0000386693

In the African landscape, steps have also been taken to recommend guidelines containing principles that recognise particular and distinctive socio-cultural nuances. Some of these initiatives include:

- AI For Africa Blueprint[12]
- African Union Resolution 473 on human and people's rights and AI[13]

The guidelines above reflect various AI principles as shown in table 1 below.

Principle	Description
Inclusive growth, sustainable development and well-being[14]	This principle emphasizes the importance of guiding AI development and use toward prosperity and beneficial outcomes for both people and the planet, highlighting its role in advancing inclusive growth, sustainable development, and well-being. It acknowledges the risks of exacerbating inequalities and biases, advocating for responsible stewardship, multidisciplinary collaboration, and public dialogue to ensure AI empowers all members of society and fosters public trust.

12 https://smartafrica.org/wp-content/uploads/2023/11/70029-eng_ai-for-africa-blueprint-min.pdf

13 https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-human-and-peoples-rights-and-art

14 https://oecd.ai/en/dashboards/ai-principles/P5

Principle	Description
Human-centred values and fairness[15]	AI development should adhere to human-centred values, including fundamental freedoms, equality, fairness, and privacy, to ensure that it aligns with human rights and promotes social justice. Implementing safeguards, human intervention, and oversight in AI systems can protect human rights, reduce discrimination, and build public trust, while measures like human rights impact assessments and ethical codes further promote fairness and value alignment.
Transparency and explainability	This principle is about transparency and responsible disclosure around AI systems to ensure that people understand when they are engaging with them and can challenge outcomes.
Robustness, security and safety	AI systems must function in a robust, secure and safe way throughout their lifetimes, and potential risks should be continually assessed and managed
Accountability	Organisations using AI should be accountable for the proper functioning of AI systems and for the respect of the other principles, based on their roles, the context, and consistent with the state of art.
Decolonisation	Decolonisation of AI involves challenging colonial perspectives and creating more inclusive, culturally varied approaches to artificial intelligence (AI). It aims to undo colonial legacies and envision alternative futures beyond dominant AI paradigms.

15 https://oecd.ai/en/dashboards/ai-principles/P6

Principle	Description
Solidarity	Solidarity as an AI principle means sharing of the prosperity created by AI, implementing mechanisms to redistribute the augmentation of productivity for all; and sharing of the burdens, making sure that AI does not increase inequality and no human is left behind. Solidarity as an AI principle also assesses the long-term implications before developing and deploying AI systems, so no groups of humans become irrelevant because of AI systems

2.2. Opportunities

The development and adoption of responsible generative AI guidelines tailored to SMEs will enhance responsible adoption of generative AI and unlock the following opportunities. There is a potential to enhance creativity and innovation through automated idea generation, prototyping, efficient data analysis for market trends, and creative content generation. Responsible integration of generative AI systems can also significantly increase efficiency and productivity. This can be through optimised resource allocation (e.g inventory management), improved data analysis and automated of repetitive tasks. Another opportunity for SMEs is that it can help to personalise customer experiences. By understanding individual customer behaviour through analysis of vast amounts of data, SMEs can tailor their services to meet specific customer needs. Integration of generative AI systems can also play a crucial role in facilitating market expansion for SMEs by providing relevant market research, predictive analytics for target identification, supply chain optimisation and language translation.

It can also assist in risk mitigation and management. Generative AI-driven tools can be developed for continuous monitoring of market conditions as well as internal and external factors that may pose risks. Such systems can predict potential challenges, identify vulnerabilities, and support crisis response planning. In essence, the responsible integration of generative AI systems can differentiate SMEs in the market, build trust with customers and stakeholders and thus can help gain competitive advantage.

3. RISKS, CHALLENGES AND MITIGATION STRATEGIES

Here we provide guidance on the risks, challenges and mitigation strategies for the responsible use of generative AI.

Engineers at Samsung's semiconductor arm used AI to help fix problems with their source code. In doing this, these employees unwittingly leaked the company's confidential data, including the source code for a new program, as well as notes from internal meetings relating to their hardware."This is because text generation AI systems retain user input data for further training. This incident led Samsung to ban the use of ChatGPT among employees.

3.1. Risks

As part of our guidelines, we aim to identify how risks emerge in various SMEs. This will involve monitoring the EU approach to risk levels. However, we envisage that the adoption of generative AI by SMEs generates several risks. Some of these risks are general risks pertinent to the use of AI, while others are particular to SMEs and the nature of their business ecosystem. The general risks of adopting generative AI might raise specific dangers such as:

- **Bias** – AI systems replicate patterns in the data they are trained upon. If this data is incomplete in some way, for example, if more is known about some parts of the population than others, then the AI systems can have different results for different groups. AI can also replicate historical patterns of unequal treatment.

- **Technology reliance and deskilling** - where people come to depend upon an AI system to do part of their job and become less able to understand when a system has gone wrong or has failed, they may take the results of AI systems as unquestionable true and never think to challenge or check them.

- **Global labour conditions**[16]**, sustainability**[17]. AI systems can be labour and resource intensive. Some have used low-paid labour to train AI models, and the process can use large amounts of energy, water and rare materials.

16 https://time.com/6247678/openai-chatgpt-kenya-workers/

17 https://www.accenture.com/il-en/blogs/consulting/making-generative-ai-green

- **Perceived or actual IP infringement, inappropriate mimicry** - generative AI are typically trained on very large datasets of images or texts. The legal status of these is not always clear. The outputs of generative AI may potentially infringe on copyright and intellectual properly.

- **Error based on poor quality training data** – sometimes, there is just not enough training data available on topic or issue to properly train an AI. This is particularly the case for rare or unusual events. This means that these areas can be more prone to error or mistakes.

- **Errors arising from misuse**[18] - AI can just be used incorrectly, or for different purposes than it was created for. Whilst generative AI seems to have a very wide range of uses, we are still learning what it is good for, and what it is not.

- **Dual use** - where AI developed for civilian applications can also be used for military applications.

The use of generative AI also poses more specific risks, such as platform risks, negative impacts on employee satisfaction / feelings of job security, and Shadow AI.

18 See, e.g., https://www.theguardian.com/technology/2024/jan/20/dpd-ai-chatbot-swears-calls-itself-useless-and-criticises-firm

3.1.1. Platform risks

Many gen AI tools offer access to the software running on the providers' servers, typically through an application programming interface (API) or a web interface. This makes it easy to start using these tools and limits the start-up cost involved in doing so. Some companies are even building business models around offering an easy to use or specialist interface on top of a generative AI platform.

However, this type of model can lead to risks for businesses that become reliant on these tools. For example, API access could be cut off, be unreliable, or the price could suddenly increase. The end user is somewhat at the mercy of the provider.

We can draw a warning from what some amazon marketplace sellers have experienced[19]. Amazon is a platform that allows small businesses to sell goods through it. It appears to have made use of internal data on what is selling well, in order to build and offer its own line of competing products. We could imagine a similar risk with generative AI providers who could see what successful business models are being built on top of their platforms, and then replicate these.

3.1.2. Negative impacts on employee satisfaction / feelings of job security

Generative AI has the potential to significantly impact employee satisfaction and feelings of job security, with both positive and negative consequences. While automation driven by generative AI

19 https://www.reuters.com/investigates/special-report/amazon-india-rigging/

can streamline processes and enhance efficiency, it may also lead to job displacement, skill redundancy, and increased workload or stress for employees. Moreover, the introduction of AI technologies can undermine employees' sense of control or autonomy over their work and raise ethical concerns related to content generation and organisational values. Additionally, the erosion of trust and morale within the workplace may occur if employees perceive management's decision to implement AI as prioritizing efficiency over human well-being.

3.1.3. 'Shadow AI'

Shadow AI is a term given to the use of AI tools by employees without the awareness of senior management or leadership, often in ways which may violate more formal policies on the use of the AI that a company may have adopted.[20] Examples would include an employee putting customer details into ChatGPT through a web browser or on their phone and asking it to generate an email to them.

3.2. Further Challenges

There are further challenges facing SMEs integrating generative AI systems. These are both global and local.

3.2.1. Global

Some of the challenges that SMEs face globally include but are not limited to difficulties of anticipating commercial opportunities

20 https://www.forbes.com/sites/delltechnologies/2023/10/31/what-is-shadow-ai-and-what-can-it-do-about-it/

accurately; high cost of implementation; lack of availability of data; a complex regulatory landscape; limited technical and human infrastructure for implementing interoperable systems; and security risks. Others include difficulties of ensuring data privacy and appropriately determining intellectual property rights. Ensuring that these systems are scalable with the growth of the business is also a critical challenge. SMEs may also often struggle to clearly quantify the return on investment for responsible integration or have employee resistance/lack of understanding of the technology. Finally, there is the challenge of ensuring this integration is environmentally friendly, since training generative AI models and inference processing have significant carbon footprints.

3.2.2. UK and Africa

In addition to the above challenges identified as global, there are also particular challenges SMEs in the UK and Africa can face. These include:

- Market Competition
  - » The presence of other enterprises working similar integration initiatives (including larger ones) with more financial resources

- Exploitation of labour
  - » The tendency to exploit cheap labour in Africa

- Localisation and Cultural Context
  - » Difficulty of localising services considering diverse cultural contexts

- Data Availability and Quality
  - » Often lack of data or quality data especially in Africa

- Access to Skilled Talent
  - » Limited human infrastructure for responsible integration

Some of the challenges highlighted in this sub section (UK and Africa) showcase a blend of contextual nuances and global relevance.

3.3. What might SMEs do wrong with generative AI?

There are irresponsible practices that some SMEs might want to experiment with. Avoiding these would be a good part of shifting towards the responsible adoption of generative AI.

- **Replacing a human-led process with generative AI output, when there is a legal responsibility to do the process.**
  - » There are some processes where a person has a duty to think about something and come to a conclusion (for example a risk assessment, health and safety assessment, or a data protection impact assessment). There's often a documentation of this process, and businesses might be tempted to use generative AI as a short-cut to create this documentation, but they would not be fulfilling the legal requirement itself.

- **Damage their brand or reputation.**
  - » Companies with brands based around authenticity, a personal touch, and human contact might experience customer push-back around their use of generative AI, which might be seen as conflicting with these values.
- **Replacing internal practices or processes when generative AI does not have the real capability to provide that process.**
  - » Businesses might be driven to reduce costs, reducing time spent on a task or even cover a lack of specialist expertise by using generative AI for tasks where it is not competent, perhaps because of misunderstanding the capabilities of the tools.
  - » For example, providing feedback to employees, creating contracts, creating internal processes, writing communications with staff etc, or any area where the output might seem plausible, but is not actually appropriate for the company or the context of use.
  - » e.g. asking ChatGPT to generate a contract for a particular customer. This will likely produce a plausible looking contract, that might be *good enough* for some circumstances, but will not reflect specialist needs of the parties, or really understand the deal being made.
- **Not checking on content generated by generative AI.**
  - » SMEs will be tempted to use Generative AI to produce marketing and promotional content. However, Generative AI is prone to 'hallucinations' - making statements of fact which are not true. This can potentially expose companies to

risks around discriminatory speech or images (e.g., racism, sexism, ablism, ageism, homophobia etc), but also to libellous or misleading content. Content generated by generative AI might also be dangerous by giving unsafe advice (e.g., DIY, health, exercise, legal advice, and many other topics).

- Similarly, Generative AI is increasingly being included in search tools, to provide summaries of information. SMEs may find this a useful way to find useful information, but 'hallucinations' in this context can include false or untrue information.

- **Providing inappropriate data to generative AI**
  - Users can provide data to generative AI tools and ask the tool to interpret, summarise or analyse this data, or to use this data in the product of some other output. This is quite a powerful tool, but SMEs need to consider if they should be sharing this data with the organisation providing the generative AI (which they might be doing if they upload it to a tool). It might be personal data, or commercially confidential.
- **Not training staff or setting expectations** about how generative AI tools should be used in the workplace.
  - Some generative AI tools are easily accessible, and employees might start using them on their own. Employers should be clear with staff about what are appropriate or inappropriate uses.
  - On the other hand, because of the way that generative AI is more accessible than earlier AI, employees can become a good source of ideas on how to make use of generative AI in the workplace.

- **Ignoring generative AI**
  - » SMEs might simply assume that generative AI is not important to them and that they do not need to think about it.
  - » They might suffer a competitive disadvantage in relation to other companies who can make responsible use of the technology.
  - » It leaves them open to the risk of 'shadow AI' - where their employees adopt generative AI tools without permission.

In February 2024 The "Glasgow Willy Wonka Experience", touted as a whimsical journey through a chocolate factory akin to the beloved Willy Wonka tale, ended up as a letdown for attendees. Despite assurances of magic, the reality fell short with a drab venue, cheap props, and an AI-generated script that failed to capture the essence of the story. Misleading advertisements using AI-generated imagery left disappointed children and fuelled internet memes. Priced at £35 per ticket, outraged parents sought refunds and involved the police, leading to the event's premature closure. The disenchanted Oompa-Loompa and the perplexing "Unknown" character only added to the confusion, highlighting the risks of relying solely on AI for event planning and the importance of human oversight and creativity.

3.4. Mitigation Strategies

To mitigate the above risks and challenges, it is important for SMEs to use responsible research and innovation (RRI) tools for effective identification of risks and responsible actions. For instance, the AREA (anticipate, reflect, engage and act) framework can help the SME to; anticipate the potential social, economic, and ethical impacts of integrating generative AI systems; reflect on the ethical implications

of deploying generative AI systems, engage with relevant stakeholders throughout the AI integration process and act by implementing clear ethical processes, actively monitor the systems' performance and impact, and act in accordance with relevant legal and regulatory requirements.

3.4.1. Create a generative AI policy

An essential initial part of the process for mitigating against the risks identified above is to create a policy on the use of generative AI. When implemented effectively, this can lead to both internal and external benefits, including setting clear expectations for employees to abide by, providing certainty and transparency to customers, and supporting regulatory compliance and adherence to ethical AI principles. Whilst the exact content will vary between different SMEs, any organisational generative AI policy should, as a minimum, clearly identify acceptable and responsible uses as well as relevant restrictions on inappropriate uses.[21] In addition, any such policy should account for potential non-compliance by establishing a procedure for reporting, investigating and resolving alleged breaches. The following subsections elaborate further on the key aspects of a generative AI policy for SMEs, the exact features of which will need to be monitored and updated as necessary, starting with the need to form responsible data governance practices.

3.4.2. Ensure responsible data governance practice

To mitigate ethical, legal, technical, and socio-cultural issues associated with the use of data in AI, SMEs need to develop responsible data governance mechanisms based on the Process, Policies and

21 See, e.g., https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai

Technology (PPT) framework. This involves developing appropriate processes through reasoned policies and agile technologies. These processes include for data management (e.g., storage, access, deletion, sharing), data protection (e.g., processing for addressing data subject rights, data protection impact assessment, identification of lawful basis for processing), ensuring data quality and diversity. These processes must have foundations in internal policies and strategies. For instance, data access policy that defines who has access to data and conditions for this access, policies on data generation and deletion as well as data sharing policies. These can as well be put into one big policy document called data governance policies. Ways to control data given to 3rd parties that are subsequently used to train AI systems can be included in this policy document. The processes and policies can be tied together with agile technology to help the preservation of privacy and confidentiality, avoidance of data leakage and compliance with legal provisions and ethical principles. These can include privacy preserving technologies (e.g., end-to-end encryption, differential privacy, homomorphic encryption, federated learning), consent management systems, data auditing and monitoring tools.

3.4.3. Ensure appropriate skill mix and training

In mitigating risks associated with generative AI, it is crucial to focus on the composition of the team's skills and the training they receive. Ensuring a balanced skill set is foundational for the responsible use of generative AI technologies. This involves assembling a team where each member's expertise complements the others, covering a broad spectrum of capabilities from technical proficiency in AI systems to ethical considerations and impact assessment.

The significance of having a well-rounded skill set cannot be overstated. Competencies are needed across various levels of AI engagement, ensuring that the team is not only adept at using AI tools but also vigilant about their implications and the ethical considerations surrounding their use.

To leverage generative AI effectively while minimizing risks and preventing harm, targeted training for employees is indispensable. This training should not only aim at enhancing their technical abilities but also at fostering an understanding of the ethical use of AI. It is about enabling the workforce to harness the potential of generative AI in a way that adds value to your enterprise, while also instituting practices that ensure responsible usage.

Encouraging responsible use among staff involves creating an environment where ethical considerations are at the forefront of AI deployment. By investing in continuous education and promoting a culture of ethical awareness, SMEs can navigate the complexities of generative AI, ensuring that their use of such technologies is both innovative and conscientious.

3.4.4. Select appropriate Large Language Model (LLM)

In selecting an appropriate LLM under responsible generative AI, it is crucial for SMEs to first understand specific business needs and use cases. Evaluating model capabilities such as language understanding, generation quality, and computational requirements, alongside considering ethical and responsible AI principles like bias mitigation, transparency, and accountability. Assessing data privacy and security measures, as well as the model's fairness and inclusivity, ensures

protection of sensitive information and promotes ethical AI practices. Scalability, cost considerations, vendor reputation, and support also play key roles in the selection process. Pilot testing and continuous monitoring allow for validation of the LLM's performance and ongoing improvement, while staying informed about evolving standards and best practices ensures ethical and compliant use of generative AI technologies within SMEs. This should include an understanding of the limitations of particular models with respect to use cases and the potential for "hallucination". SMEs may also wish to consider the use of open-source models with published hyperparameters, and models trained on open data, enabling knowledge sharing across industry and fostering a culture of transparency in development with benefits to security and user privacy.

3.4.5. Ensure transparency

In technology integration, transparency is crucial for building trust, meeting ethical standards, and addressing concerns. SMEs integrating generative AI systems need to prioritise AI models that provide clear explanations for their decisions; models that are interpretable or that are inherently transparent. It is also important to maintain comprehensive documentation of the AI design and development processes wherever they are involved; detailing data sources, pre-processing steps, model architectures and parameter tuning, or their decisions; usingmodels that are interpretable. It is also important to maintain comprehensive documentation of the AI design and development process which can be made accessible to relevant stakeholders. Ensuring transparency can also include considering open-source practices where applicable and conducting regular audits of the systems' algorithms and decision-making processes.

3.4.6. Consider unintended impacts

The use of generative AI can have unintended effects that negatively impact upon a company's reputation. Consider two: bias, and sustainability challenges.

Bias. The value of AI lies in its ability to process vast amounts of data. However, this comes with the risk of perpetuating existing biases. Small enterprises may lack the resources to procure or make independent analyses of large, diverse datasets, leading to AI systems that may not be trained on representative data, or to a dependence upon external processing systems. This can result in AI models that inadvertently discriminate against certain groups, particularly in hiring, lending, and customer service applications. The consequences of AI bias for small businesses are not merely ethical but also practical. Biased AI systems can alienate segments of the customer base, damage brand reputation, and lead to legal challenges. Moreover, these biases can skew market research and business insights, leading to poor strategic decisions. Small enterprises, with their limited buffers against market fluctuations, are particularly vulnerable to such missteps.

Sustainability. Sustainable AI presents a special challenge for small businesses that may wish to claim sustainability credentials. The development and maintenance of AI systems can be costly, and the expertise required to ensure they are sustainable may be beyond the reach of many small businesses. Moreover, the energy consumption associated with running AI algorithms is a significant concern, as it can contribute to a business's carbon footprint. Small enterprises must balance the benefits of AI with the need to minimize environmental impact, which can be a delicate and complex task.

3.4.7. Consider how much you want to reply upon generative AI for business-critical functions

The involvement of generative AI in business processes creates a potential for increased operational complexity. While AI systems promise to streamline processes, the reality is that they often require significant management and oversight. Small businesses may find themselves grappling with the intricacies of AI integration, from data management to system maintenance, which can be both time-consuming and costly.

Another risk is that of alienating customers. As AI becomes more prevalent in customer service, there is a possibility that the personal touch, which is often a unique selling point for small businesses, will be lost. Customers who value human interaction may feel disconnected from a business that leans too heavily on automated systems, potentially leading to a loss of loyalty and a reduction in revenue.

3.4.8. Consider intellectual property

For enterprises using or looking to use generative AI, a core area of concern is the potential risk of liability arising from infringement with protected intellectual property (IP) rights, such as copyright.[22] Both initial training and more context specific fine-tuning of generative AI models requires a significant volume of data, a common source for which is those data that can be extracted from websites and other online environments through web-scraping and other data harvesting techniques. Whilst some of these sources are likely to be

22 See, e.g., https://www.acrolinx.com/wp-content/uploads/2023/08/Acrolinx_Generative-AI_Report_FNL.pdf

made available as open access, others may subject to more restrictive licensing conditions for reuse, such as that users request permission and/or pay a fee. Indeed, even freely accessible sources of data may be restricted in this way. As such, SMEs need to actively consider IP in their use of generative AI and take steps to mitigate against the risks.

For SMEs involved in training or fine-tuning of generative AI models, whether developed internally or, perhaps more likely, where provided by a developer via an application programming interface (API), the provenance of the dataset(s) and any conditions (e.g., licensing and compensation) attached to usage should be assessed to promote transparency (see above) and ensure full compliance with IP law. If the data used and/or the generative AI model itself is sourced from a developer or another provider, SMEs will need to ensure they carry out sufficient due diligence. In particular, consideration should be given to entering into contractual arrangements on the basis of assurances around compatibility with IP law and protection against losses from a failure to do so. Some developers, especially larger organisations, are already offering indemnities against legal and financial losses. Whilst also facing legal action from the New York Times for alleged copyright infringement, Open AI, for instance, has established a "Copyright Shield" for all "generally available features of ChatGPT Enterprise", through which it pledges to "pay the costs incurred" for customers facing legal action over copyright infringement.[23] However, as in this example, these indemnity clauses are unlikely to be broadly applicable and, as such, may offer only limited protection for SMEs against the risk of third-party IP infringement arising from the use of generative AI.

23 https://openai.com/blog/new-models-and-developer-products-announced-at-devday

3.4.9. Make the best use of support from government and other sources

There are some sources of support available to SMEs that may be able to help them explore or implement generative AI. This can include financial support from governments, funding bodies, universities, but also networks and industry bodies.

- The UK government provides tax relief on corporation tax for Research and Development (R&D) projects, allowing companies to claim back some of the costs of innovative projects in science and technology. Not all uses of generative AI will qualify, but if a project does qualify, then this can be a way of reducing the costs of responsible generative AI deployment. To qualify, a project needs to look for an advance in a field related to your company's trade, have overcome a scientific or technological uncertainty, and do something that could not be easily worked out by a professional in the field. More guidance is available from HMRC.[24]

- The Innovate UK Bridge AI programme is intended to support UK business through the responsible adoption of AI in priority sectors, such as transport, construction, agriculture and food production, and creative industries. The programme offers funding and support to help innovators assess and implement trusted AI solutions, connect with AI experts, and elevate their AI leadership skills.[25] Support from the programme includes:

24 https://www.gov.uk/guidance/corporation-tax-research-and-development-rd-relief

25 https://iuk.ktn-uk.org/programme/bridgeai/ open support opportunities are advertised here: https://iuk.ktn-uk.org/opportunities/?_sft_areas=bridgeai

- Access to the Hartree Centre Training portal of training and upskilling in AI. [26]
- Free training courses from the Turing institute, developed by Trilateral Research on operationalising ethics in AI.[27]
- This included a now closed funding competition for UK organisations to support the development and adoption of AI and machine learning solutions.[28]

- There are three regional SME engagement hubs in Cardiff, North East England and Northern Ireland providing local support from organisations with expertise in data science and AI to help SMEs.[29]

- Schemes like the AI Futures Grant provide support to SMEs to meet the relocation costs of early to mid-career AI researchers and engineers with exceptional promise moving to the UK.[30]

26 https://iuk.ktn-uk.org/opportunities/bridgeai-access-to-hartree-centre-training-portal/

27 https://www.turing.ac.uk/courses/operationalising-ethics-ai-intermediate

28 https://apply-for-innovation-funding.service.gov.uk/competition/1714/overview/c85b3edc-0416-4e92-a5d3-014fe8aab148#eligibility

29 https://www.hartree.stfc.ac.uk/sme-hubs/

30 https://www.gov.uk/guidance/corporation-tax-research-and-development-rd-relief

3.4.10. Future Development Opportunities

- The SEO market is likely to be disrupted in the near future.[31]

- As SMEs navigate this evolving landscape, it is crucial to monitor SEO metrics closely to gauge any significant shifts resulting from the increasing prevalence of Generative AI. Ongoing observation and adaptation will be key to staying ahead in such a dynamic environment. For example, as the popularity of chat-based interfaces continues to rise, it's essential for SMEs to reassess their search marketing strategies and how they align with these evolving trends. For instance, if your approach has centred on creating content-rich pages, like blogs, to attract visitors through search queries, you may need to consider the potential impact of chatbots autonomously providing answers to user inquiries. Conversely, if your objective is to showcase specific offerings, such as upcoming events or classes, the influence of chatbots might be less pronounced.

- Experimentation and continual measurement - to exploit future development opportunities while mitigating risks SMEs can adopt four different layers of potential mitigations: at the **model level**, at the **safety system level**, at the **application level**, and at the **user level**. Adopting this strategy will enhance an iterative and layered approach to identify and mitigate risks while maximising future development opportunities tailored towards individual needs and staying ahead in a dynamic generative AI adoption environment.

31 https://www.forbes.com/sites/nishatalagala/2023/03/25/chatgpt-is-seoas-we-know-itdead/?sh=12d6caaa395e

- As Generative AI has the potential to revolutionize content creation by generating high-quality images, videos, and graphics, and by creating dynamic content tailored to an individual's preferences. This can help businesses stand out in a crowded marketplace and optimize engagement and conversion rates.

> *In 2023, a New Zealand-based supermarket Pak 'n' Save created a meal planner app (Savey meal-bot) designed to help households save money and reduce food waste by suggesting meals complete with a full recipe and directions. A user shared on X (formerly Twitter) how this app suggested a recipe for deadly chlorine gas.*

3.4.11. Ensure Human Oversight

The concept of 'human-in-the-loop' (HITL) is a term in Innovation that refers to the continuous involvement of human oversight in the design, development, and deployment of AI systems. Whilst acknowledging that SMEs often do not have sufficient resources for dedicated AI experts on the team, it is critical that human oversight ensure that people with knowledge can monitor relevant stages of design and implementation. HITL can help the SME address ethical considerations, comply with regulations, become transparent and proactively detect errors. This can be achieved through;

- A robust review and approval of all recommendations made by AI systems before implementation. The team's judgement and experiences are thus leveraged on to ensure that final decisions align with business goals and ethical considerations.

- Identification and handling of complex or unexpected situations that AI systems are not built to address.

- Clear establishment of roles and responsibilities within the organisation regarding who makes the final decision between the AI and humans. It is also essential for SMEs to maintain open communication with human experts and AI designers, providing feedback on performance for improvement of the system.

4. AI LIFECYCLE POLICIES AND APPROACHES

As a fundamental aspect of the Responsible Generative AI for SMEs guidelines, our aim is to seamlessly incorporate our recommendations throughout the various stages of AI adoption by SMEs. This entails integrating responsible generative AI principles, recommendations, and practices within the following stages of the AI adoption lifecycle:

4.1. Lifecycle

The AI Lifecycle encompasses the stages in which an AI solution is conceived, developed, deployed and utilized.

- **Sourcing AI** is the first stage of the AI lifecycle and involves identifying and acquiring the necessary resources to develop and implement an AI system. This includes identifying the appropriate data, algorithms, and computing power.

Responsible AI can be integrated into this phase by:

  - » Ensuring that the data used to train the AI system is **high-quality, unbiased, and representative**.
  - » Ensuring that the algorithms used to develop the AI system are **transparent, explainable, and fair**.
  - » Ensuring that the computing power used to develop and implement the AI system is **sustainable**.

- **Developing AI** is the second stage of the AI lifecycle and involves designing and building the AI system. This includes developing the algorithms, training the model, and testing the system.

Responsible AI can be integrated into this phase by:

- Ensuring that the algorithms are designed to **meet the specific needs** of the organization.
- Ensuring that the model is **trained on a representative dataset**.
- Ensuring that the system is **tested thoroughly to identify and mitigate potential risks**.

- **Implementing AI** is the third stage of the AI lifecycle and involves deploying the AI system into the real world. This includes integrating the system into existing processes, training users on how to use the system, and monitoring the system's performance.

Responsible AI can be integrated into this phase by:

- Ensuring that the system is **implemented in a way that minimizes risks**.
- Ensuring that the system is **monitored closely to identify and mitigate any unintended consequences**.
- Providing **training for users on how to use the system responsibly**.

- **Deploying AI** is the fourth and final stage of the AI lifecycle and involves making the AI system available to users. This includes releasing the system to the public or specific user, managing the system's availability, and providing support to users.

Responsible AI can be integrated into this phase by:

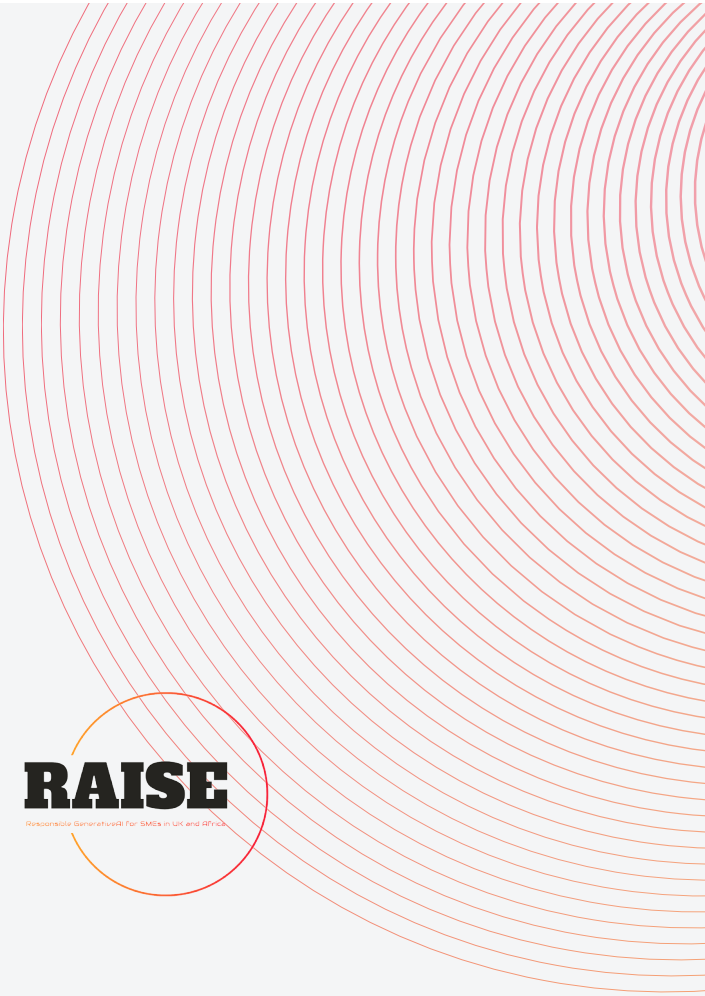
- Ensuring that the system is **deployed in a way that minimizes risks**.
- Ensuring that the system is **monitored closely to identify and mitigate any unintended consequences**.
- Providing **training for users on how to use the system responsibly**.
- Ensuring that the system is **designed to be fair and equitable**.
- Ensuring that the system is **transparent and accountable**.

We believe that integrating our responsible guidelines and leveraging the AI lifecycle and approaches across sourcing, developing, implementing, and deploying stages can help SMEs harness the benefits of AI while mitigating risks and maximizing return on investment.

5. SUMMARY

The RAISE Guidelines provide guidance on the responsible use of generative AI for small and medium-sized enterprises (SMEs) in the UK and Africa. These guidelines aim to gather feedback on content and representation, as well as to assist SMEs in avoiding potential pitfalls associated with generative AI use.

Key findings of the guidelines highlight both the risks and mitigation strategies associated with generative AI. Risks include platform risks, negative impacts on employee job satisfaction and feelings of job security, and "shadow AI." Mitigation strategies include establishing a generative AI policy, ensuring sound data governance practices, and acknowledging the additional effort and time required by SMEs to implement responsible AI practices. The guidelines also outline four key steps to integrate responsible AI into the AI lifecycle: sourcing AI, developing AI, implementing AI, and deploying AI. In the deployment phase, responsible AI involves minimizing risks, monitoring the system for unintended consequences, providing user training, ensuring fairness and equity, and guaranteeing transparency and accountability.


RAISE
Responsible Innovation for SMEs in UK and Africa