AI for Social Value

A Practical Guide for Local Authorities
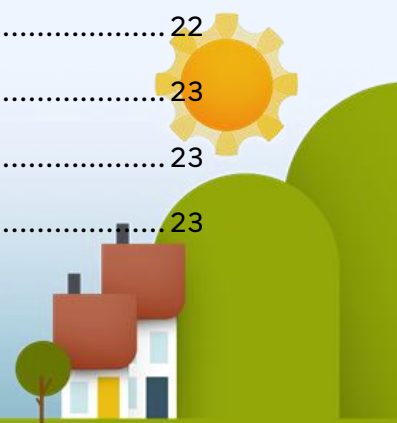
Maddie Kortenaar

www.socialvalueengine.com

Contents

Introduction

Artificial Intelligence (AI) has rapidly integrated into our personal and professional lives, with Generative AI (GenAI) tools like ChatGPT showcasing both its vast potential and significant risks.

In my professional network, creative professionals are losing jobs to AI-generated content, and the IT job market is stagnant with wages not growing. Personally, AI has been invaluable—for example, I used it to practice interviews and refine my CV when my former employer closed, allowing me to focus on enhancing my skills. This experience highlighted the complexities of incorporating AI into daily life and work.

Today, I regularly use various AI tools, understanding their strengths and limitations. I've learned to identify AI-generated work, such as repetitive phrases in text from ChatGPT or unrealistic elements in AI-created images. My work with the Social Value Engine has deepened my awareness of AI's societal implications, emphasising the need to harness its benefits while mitigating risks to ensure it serves everyone fairly.

This guide aims to introduce AI to professionals in social value, particularly within local authorities. It explores how AI can be leveraged for public good, addressing economic transformations, ethical challenges, and community impacts. We will discuss the importance of open-source tools, data ownership, and equitable distribution of AI's advantages. Ultimately, the goal is to shape AI in ways that enhance lives and promote well-being, avoiding the reinforcement of existing inequalities.

I hope this guide will enhance understanding of the intersection between AI and social value, helping us build communities where AI truly benefits society.

Maddie Kortenaar

Note: For simplicity, I use 'AI' to refer to both traditional and generative AI. I have also chosen not to delve too deeply into the technical aspects of AI, data ownership, large language models, and data provenance, as there are more specialised works available on these topics.

Chapter 1: The AI revolution and economic shifts

Integrating artificial intelligence into our daily routines and workplaces has been nothing short of transformative. In just a few years, AI has progressed from the pages of science fiction to become a practical asset employed in every corner of industry—from creative design to financial services.

This rapid advance in AI is reshaping the professional landscape, yielding developments that are both thrilling and unsettling. On one hand, AI offers unprecedented opportunities to drive innovation and boost efficiency. On the other, it is disrupting established roles and compelling us to reassess the very nature of work.

In this chapter, we examine the current extent of AI adoption and consider its impact on the job market, covering changes to wages, productivity, and economic equity. We also explore how best to manage these shifts so that the advantages of AI are shared as widely as possible. Our aim is not to position AI as either villain or saviour, but rather to understand its influence on our evolving economy and to consider how we can harness its capabilities while tackling the challenges it brings. By doing so, we can pursue a future where AI strengthens our communities instead of dividing them.

The Current State of AI Adoption

Artificial intelligence has rapidly evolved from a niche technology into a ubiquitous presence across countless industries. Businesses are automating workflows, improving service delivery, and even transforming everyday human–technology interactions. Organisations seeking to measure and report on social impact, for instance, are increasingly turning to AI-powered tools—work that was once labour-intensive and time-consuming.

However, AI's reach goes far beyond specialist applications:

- **Healthcare:** Medical practitioners use AI for tasks such as diagnostic imaging and drug discovery, accelerating and enhancing patient care.
- **Finance:** Banks and other financial institutions rely on AI for fraud detection, algorithmic trading, and personalised advice. The ability to process vast amounts of financial data quickly and accurately makes AI indispensable in this sector.
- **Retail:** Online retailers routinely employ AI-driven recommendation engines and chatbots to deliver tailored customer experiences. These systems learn continuously from user behaviour to refine their services.
- **Manufacturing:** AI plays a key role in optimising supply chains, predicting equipment failures, and improving quality control.
- **Creative Industries:** Perhaps most strikingly, AI is now generating original content— writing articles, creating images, and even composing music. This prompts new questions about creativity and authorship.

The pace of AI adoption shows no sign of slowing. As tools become more accessible and user-friendly, small businesses and non-profit organisations are also starting to integrate AI into their operations. This "democratisation" of AI is exciting, but it raises important questions about the future of work and the skills that will be needed in an AI-driven economy.

Economic Implications

The rapid adoption of AI is reshaping industries, creating opportunities while posing challenges to the workforce. According to a study by the IMF in 2024, up to 60% of roles in advanced economies are susceptible to automation, with high- and low-skilled jobs equally affected.

The integration of AI into our economy is producing wide-ranging effects—some are already evident, while others are only just coming into view. In many sectors, especially within IT and creative industries, wage growth has slowed significantly, a trend that appears to coincide with the expanding use of AI tools.

For example, entry-level programming positions—once a dependable route to a well-paying career—are now under pressure from AI coding assistants capable of handling numerous basic development tasks. Likewise, in creative fields, the speed and low cost of AI-generated content are driving down freelance rates.

On a broader economic scale, AI has the capacity to significantly enhance productivity. By automating routine tasks and augmenting human capabilities, AI can enable organisations to achieve more with fewer resources. This efficiency may lead to economic growth and, in principle, greater overall prosperity.

However, a pressing question arises: who will benefit from this increased productivity and potential economic growth? Will it primarily serve shareholders and executives, or will it be distributed more widely across the workforce and throughout society?

This consideration leads to one of the most consequential economic implications of AI: its potential to deepen existing inequalities. As more roles become automated and the demand for advanced technical competencies grows, we risk perpetuating a two-tiered economy.

On one side, individuals with the skills to work alongside AI may see their productivity—and, by extension, their earnings—surge. On the other side, workers whose skills are devalued by automation may struggle to secure well-paid employment.

Crucially, this is not solely about mastering cutting-edge technology. The modern economy places a growing premium on uniquely human attributes

such as emotional intelligence, creativity, and the ability to solve complex problems. Those who can combine these qualities with a solid grasp of AI tools are likely to thrive in the evolving landscape.

Can We Get It Right This Time?

The rapid growth of AI mirrors the transformative impact of the internet, which redefined industries and created new opportunities. While early adopters thrived, companies like Kodak and Blockbuster declined due to failure to adapt. In today's AI-driven world, avoiding "AI illiteracy" is essential, as it risks becoming a significant barrier to employment and economic participation.

Although the internet brought unprecedented connectivity and access to information, it also introduced a host of complex issues we are still grappling with today. These include the exploitation of personal data, the spread of biased news sources, and business models driven primarily by profit—often at the expense of individual and community well-being.

AI offers a chance to address challenges proactively by prioritising

ethical AI systems that respect privacy, ensuring accessible AI education for all, balancing profit with public good through social value principles, and fostering critical thinking to navigate an AI-driven world confidently.

Opportunities for a More Connected Society

AI presents challenges but also the potential to build more connected communities. By automating repetitive tasks responsibly, it can free people to focus on meaningful work and foster human skills like empathy, creativity, and problem-solving. AI also presents innovative ways to improve education and care services:

- **Personalised Education**

  AI-driven learning platforms can adapt to each student's individual needs, resulting in more engaging and effective educational experiences.

- **Enhanced Healthcare**

  From assisting in diagnoses and treatment plans to monitoring patient well-being, AI can streamline healthcare
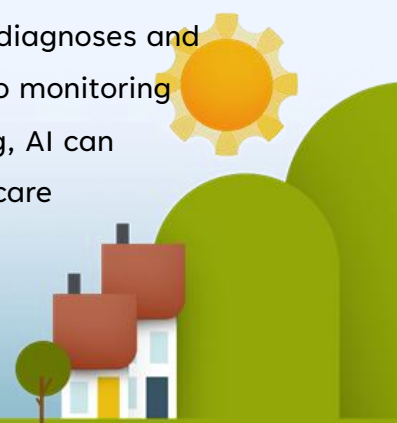
processes, enhance outcomes, and improve access to quality care.

When AI is used to boost human capabilities—rather than diminish them—it can help preserve the compassion and personal touch essential to education, healthcare, and other areas that rely on meaningful human interaction.

Fostering Community Connections

As AI increasingly handles routine tasks, it may paradoxically create more opportunities for genuine human interaction. Freed from certain administrative or repetitive duties, people might choose to devote more time to volunteering and local initiatives, potentially leading to a resurgence of community engagement.

AI can also directly support these connections. For instance, AI-powered platforms could match volunteers with local roles that suit their skills and interests, or facilitate skill-sharing between neighbours. By helping people locate and engage with one another, such systems could reinforce community bonds.

Moreover, AI holds promise for tackling complex societal challenges. Advanced modelling tools can help us better understand and address issues like climate change, while predictive analytics can guide targeted interventions in social services. For example, AI systems might spot early indicators of social isolation or mental health concerns, enabling timely support. They could also help optimise the allocation of resources for community programmes, ensuring that assistance reaches the people who need it most.

By centring AI development on communal well-being, we can create an AI-enhanced society that strengthens our shared humanity, rather than undermining it.

Chapter 2: AI Bias and Societal Reflection

AI bias refers to systematic, repeatable errors in AI systems that lead to unfair outcomes. These biases can manifest in various ways—such as facial recognition systems that struggle with darker skin tones or language models that reinforce gender stereotypes. In many cases, the data used to train the AI lies at the heart of the issue. If the data is unrepresentative of the entire population or contains historical biases, the AI is likely to mirror and sometimes amplify these biases.

For example, large language models like GPT are trained on data taken from online forums such as Reddit, which capture a narrow slice of public opinion. Another instance could be an AI system designed to streamline hiring in a historically male-dominated field. Even if gender is never explicitly considered, the system might learn to prefer male applicants because it is basing its decisions on biased historical outcomes.

Design choices can also create bias if the team developing an AI system lacks diversity or fails to account for different perspectives. Such biases can be subtle, and they often remain hidden until patterns of unfair outcomes emerge over time. This makes ongoing monitoring and evaluation of AI systems essential.

Types of AI Bias

AI bias arises when systems reflect or amplify inequalities present in their training data. For instance, facial recognition tools have struggled with darker skin tones due to underrepresentation in datasets, while hiring algorithms have perpetuated gender biases in traditionally male-dominated fields. These issues highlight the importance of diverse, representative data and rigorous testing during development.

Societal Impacts of AI Bias

For public sector bodies, AI bias poses significant risks to service delivery and community well-being. Recognising and addressing these issues is crucial to ensure fair and effective governance. Below are some key areas where biased AI systems can have far-reaching consequences:

1. **Social Care and Support Services**

   AI-driven assessments for allocating resources or support could inadvertently marginalise certain groups if the underlying data or algorithms are not properly calibrated. Vulnerable populations—such as older adults or disabled individuals—may be disproportionately affected, leading to unequal access to essential services.

2. **Housing and Urban Planning**

   AI-based models for urban development or housing allocation might perpetuate existing patterns of segregation or gentrification if they rely on historically biased data. This can exacerbate inequalities and fuel tensions within communities.

3. **Local Economic Development**

   Tools used to guide economic strategies or direct business support may favour particular demographics or areas, potentially widening economic disparities.

4. **Public Transportation**

   Algorithmic planning tools could inadvertently favour specific routes or districts, restricting mobility and access to opportunities for under-served groups.

5. **Community Safety**

   Predictive policing algorithms and CCTV analytics, if based on biased

historical data, may disproportionately target certain neighbourhoods or demographics, eroding trust and fairness.

6. **Citizen Engagement**

AI-driven chatbots or automated response systems that fail to recognise certain dialects or queries could discourage or exclude individuals seeking information or assistance.

Such biases can undermine public sector bodies' mission to serve all community members equitably, weaken public trust, and reinforce existing inequalities. As these organisations increasingly rely on data-driven decision-making, there is a real risk of creating a cycle in which biased AI leads to unfair resource allocation and service provision, further entrenching social disparities.

Strategies for Mitigating AI Bias in Local Governance

Addressing AI bias is essential for local authorities seeking to deliver fair and effective services. Below are key strategies to help minimise bias and

promote equitable outcomes. Remember that tackling AI bias is an ongoing process, requiring constant vigilance and adaptation as both technology and community needs evolve.

1. **Diverse and Representative Data**

Ensure the datasets used to train AI systems include the full spectrum of community demographics. This may involve supplementing existing data with targeted collection efforts to capture the experiences of underrepresented groups.

2. **Regular Audits and Impact Assessments**

Conduct frequent evaluations of AI tools to detect and address potential biases. In addition, perform algorithmic impact assessments before deploying new AI technologies, considering their effects on various community segments.

3. **Transparency and Explainability**

Use AI systems that clearly document how decisions are reached. Openness about how

algorithms function supports scrutiny, fosters public trust, and helps identify hidden biases.

4. Community Engagement

Involve diverse stakeholders in the development and testing of AI solutions. Community members can offer insights that may not be evident to the design team, helping to spot and mitigate biases early.

5. Cross-Departmental Collaboration

Encourage cooperation between IT teams and frontline service departments. Those with deep knowledge of local needs and challenges should play a role in AI planning and implementation.

6. Ongoing Monitoring and Adjustment

Continuously measure the outcomes of AI-based decisions. Be prepared to recalibrate systems if disparities or unintended consequences emerge.

7. Diverse Development Teams

Whether AI solutions are developed in-house or procured from vendors, prioritise team diversity. A range of perspectives can help identify and correct biases during the design phase.

8. Clear Governance Structures

Define clear lines of responsibility for AI usage. Assign accountability to specific roles or committees to ensure prompt action when issues arise.

9. Training and Awareness

Offer training sessions to help staff recognise and reduce AI bias within their departments. Building understanding among employees is a vital step toward equitable outcomes.

10. Ethical Guidelines

Develop and follow robust ethical frameworks for AI usage, prioritising values such as fairness, accountability, and inclusivity. Embed these guidelines into every stage of AI design and deployment.

11. Open Source and Shared Learning

Consider open-source AI solutions and participate in networks with other local authorities. Sharing knowledge and experiences can accelerate collective efforts to address bias effectively.

Human Oversight and Social Value in AI Implementation

While technological solutions play a critical role in addressing AI bias, human oversight and a commitment to social value are equally important. For public sector bodies, this means integrating AI within a broader framework of ethical governance and community benefit.

Human judgement offers the contextual understanding, ethical considerations, and common sense that AI may lack. In sensitive areas such as social care assessments or housing allocations, final decisions should always involve a human perspective.

Incorporating Social Value

When deploying AI for decision-making, or implementing new AI-powered solutions, public sector bodies should weigh efficiency gains against wider community outcomes. This approach emphasises the creation of social value alongside financial or operational benefits. In this context, organisations may wish to consider the following questions:

1. **Community Cohesion**

   How does the AI-enabled solution affect different community groups? Does it foster inclusivity or risk marginalising certain demographics?

2. **Local Economic Benefit**

   Does the system create opportunities for local residents and businesses, or does it limit them?

3. **Wellbeing**

   How does the AI solution impact overall community wellbeing, including physical, mental, and social health?

4. **Environmental Impact**

   Does it align with sustainability objectives or undermine them?

5. **Community Engagement**

   Does the AI-driven process encourage greater involvement in local governance or inadvertently discourage public participation?

Robust ethical guidelines for AI are essential. These frameworks should be developed collaboratively, involving not only technical experts but also ethicists, community representatives, and service users. Such an inclusive approach helps guarantee that AI systems reflect shared values, uphold accountability, and contribute meaningfully to community well-being.

Transparency and Accountability

Public sector bodies should maintain a high standard of transparency when implementing AI tools and processes. They should cultivate a culture of continuous learning, remaining agile in adapting AI strategies as new issues and possibilities arise.

By integrating human oversight and social value considerations into their AI approaches, organisations can ensure that powerful technologies genuinely serve the diverse needs of the community. In doing so, they not only mitigate bias but also build public trust, reinforcing their fundamental mission of delivering equitable and effective services for all residents.

Chapter 3: Human Interactions in Social Value Assessment

As artificial intelligence continues to reshape local governance, we must not lose sight of an irreplaceable component in creating and measuring social value: human interaction. While AI can analyse vast datasets and pinpoint patterns, it is human insight, empathy, and local knowledge that often provide the broader context for understanding true community needs and the real-world impacts of policy decisions.

For public sector bodies, finding the right balance between utilising AI's capabilities and preserving meaningful human engagement is crucial. Communities are intricate, diverse ecosystems where quantitative data alone cannot capture the full story. Genuine human interaction offers the nuanced perspective necessary for effective, responsive local governance.

In this chapter, we will explore why human involvement remains essential in social value assessment, even as AI becomes more integrated into our decision-making processes. We will discuss how qualitative insights gained from human-led approaches can highlight aspects that AI might overlook, how community engagement should guide the development of AI solutions, and practical strategies to combine human expertise with AI-driven tools to maximise social value for all.

The Unique Strengths of Human Interaction

While AI excels at analysing large datasets and identifying patterns, people bring qualities to social value assessment that technology simply cannot replicate. Through empathy and emotional intelligence, human staff can discern subtle cues and contextual details, deepening their understanding of community needs and ensuring that trust-building remains at the heart of governance. Their capacity for flexibility and adaptability enables them to respond swiftly to unexpected information, while their local knowledge helps to interpret nuances and avoid oversights that AI might miss.

Human interaction also enriches social value assessment with qualitative insights drawn from stories, opinions, and personal experiences. Long-term relationships with residents and community organisations offer invaluable perspectives on shifting challenges and emerging opportunities. Equally important, people can exercise complex ethical judgement, creativity, and cultural sensitivity—essential traits for navigating diverse social contexts. By blending these strengths with AI's

analytical power, public sector bodies can develop a more complete, responsive, and inclusive approach to serving their communities.

Integrating Human Insight with AI Capabilities

While human interaction is vital for capturing nuances and context, AI can significantly enhance social value assessments by sifting through large datasets, identifying patterns, and automating routine tasks. The key is to balance and integrate the strengths of both human insight and AI. Below are several approaches public sector bodies can use:

1. **AI-Assisted Data Analysis**

   Use AI to process large volumes of data swiftly, highlighting trends and patterns. Human analysts can then interpret these findings, adding local knowledge and contextual understanding to shape more meaningful conclusions.

2. **Sentiment Analysis Tools**

   Leverage AI to analyse text from surveys, social media, and other sources to gauge public
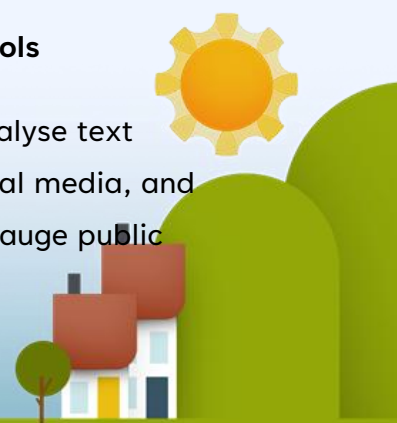
sentiment. Staff can employ these insights as a springboard for deeper, more personal community engagement.

3. Predictive Modelling

Deploy AI to model potential outcomes for various initiatives. Human decision-makers can then weigh these projections against on-the-ground realities, considering factors the AI might not capture—such as historical context or emerging social concerns.

4. Chatbots for Initial Engagement

AI-driven chatbots can handle straightforward queries and gather basic information. Complex issues or nuanced situations are then escalated to human staff, ensuring empathy and deeper understanding where needed.

5. AI-Enhanced Surveys

AI tools can help design and analyse surveys for efficiency. Complement these tools with face-to-face interviews or focus groups conducted by human staff to gain richer, qualitative insights.

6. AI-Supported Resource Allocation

Algorithms can suggest how to allocate resources based on data analysis. Human leaders can refine these recommendations, drawing on their understanding of local priorities and specific community needs.

7. Bias Detection

Use AI to spot potential biases in decision-making processes or data sets. Human teams then investigate these findings in detail and develop targeted strategies to reduce inequalities.

8. Continuous Learning Loop

Feed human feedback and real-world outcomes back into AI models to refine their accuracy and relevance, creating an ongoing cycle of improvement.

9. Ethics Committees

Establish human-led ethics committees to oversee AI usage in social value assessments. These committees ensure that AI tools are deployed in ways that align with community values, fairness, and inclusivity.

Challenges and Considerations in Balancing Human Interaction and AI

Although combining AI capabilities with human insight offers significant benefits for public sector bodies, a range of potential pitfalls must be managed responsibly. Data privacy and consent remain paramount, requiring compliance with regulations such as GDPR and ongoing attention to informed consent. Meanwhile, insufficient staff skills can limit AI's usefulness, necessitating effective training programmes. Overreliance on AI carries its own dangers, as it can crowd out vital community perspectives and qualitative data, highlighting the importance of clear guidelines on when human judgement should take precedence.

In addition, equity and transparency need to be safeguarded by offering inclusive engagement options and ensuring AI systems are explainable to both staff and the public. Resistance to change is natural, but it can be mitigated through open dialogue and balanced communication about AI's

benefits and limitations. Cost and resource implications should be weighed against the need to preserve human-led services, particularly in areas where empathy and personal connection are crucial. Ongoing ethical reviews can help maintain fairness and accountability, especially given AI's struggles with cultural nuances, rapidly shifting community dynamics, and the risk of perpetuating existing biases. By addressing these challenges proactively, public sector bodies can responsibly harness AI's potential without sacrificing essential human values and insights.

Avoiding AI 'Echo Chambers'

As AI becomes more integral to our work, it's important to avoid situations where AI ends up primarily interacting with itself—a phenomenon known as an "AI echo chamber." This occurs when AI systems learn mainly from other AI systems or outdated information, rather than reflecting real, current community experiences.

For instance, imagine using one AI system to draft emails about community projects, another AI to analyse these emails for public
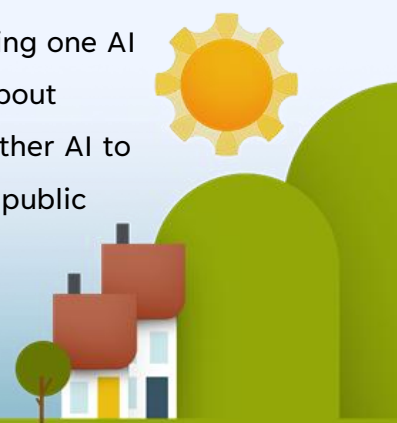
sentiment, and a third AI to plan future initiatives based on that sentiment. In this scenario, there is a risk of creating a closed loop where AI keeps informing AI, with no meaningful human input. It becomes a digital version of "Chinese whispers" where inaccuracies may be amplified through each stage.

Without human oversight at key points, decisions could be made using information that has become increasingly detached from what is actually happening in the community. In other words, we could lose touch with people's real needs, concerns, and experiences—a result that runs counter to the very purpose of implementing AI in public service.

Best Practices for Maintaining Effective Human Interaction in an AI-Enhanced Environment

As public sector bodies continue to adopt AI tools, it's vital to preserve and even strengthen human engagement. Below are some key practices to ensure meaningful human interactions remain at the heart of public service:

1. **Prioritise Face-to-Face Engagement**

   While AI can automate many processes, continue to hold in-person meetings, community forums, and local events. These personal interactions build trust and reveal insights that AI cannot capture on its own.

2. **Train Staff in 'AI–Human Collaboration'**

   Provide targeted training to show employees how to use AI-generated insights effectively, treating AI as a tool that supports—rather than replaces—their direct engagement with community members.

3. **Emphasise Active Listening**

   Encourage staff to practise active listening, especially when balancing AI-driven data with personal stories and accounts. This helps ensure that residents feel heard and respected.

4. **Maintain Multiple Communication Channels**

   Offer a variety of ways for residents to interact—online platforms, phone lines, and face-to-face meetings—so that everyone can

choose the channel that suits their needs.

5. Regular Community Feedback Sessions

Host frequent gatherings where community members can comment on AI-influenced decisions and processes, ensuring their perspectives continue to shape local initiatives.

6. Cultivate Emotional Intelligence

Invest in emotional intelligence training for staff. This helps them better understand and respond to the needs of diverse community groups—an area where AI on its own often falls short.

7. Personalise Communication

Use AI insights to tailor your communication strategies but keep the actual conversations and engagements human-led to preserve authenticity.

8. Transparent Decision-Making

Clearly communicate how decisions are reached, explaining the roles played by both AI analytics and human judgement. This transparency builds public trust in new technologies.

9. Continuous Human Oversight

Put processes in place for regular human review of AI-generated reports and recommendations. Human oversight ensures that local context and ethical considerations are not overlooked.

10. Storytelling and Narrative

Encourage staff to use storytelling when sharing information. Personal narratives give emotional depth to AI data and convey the real impact on people's lives.

11. Empathy-Driven Design

Incorporate empathy and user experience into every stage of AI system design. Make sure that technology enhances, rather than diminishes, meaningful human connections.

12. Ethical Reflection Sessions

Schedule regular team discussions focused on ethical implications of AI use, where staff can examine real scenarios and challenges to stay aligned with core community values.

Chapter 4:
Improving AI Models with Credible Social Value Data

Effective decision-making and service delivery rely on high-quality data, and this is especially true in the context of social value assessments, where outcomes can profoundly impact community well-being. Ensuring that AI models have access to credible, comprehensive social value data is therefore critical.

In this chapter, we explore the ways in which public sector bodies can enhance their AI models by incorporating reliable social value data. We will consider how to identify and use credible sources and discuss methods for collecting, validating, and integrating this data into AI systems.

The Importance of Credible Social Value Data in AI Models

Reliable data is the foundation of effective AI tools. Accurate, diverse datasets ensure fair resource allocation and transparency in decision-making. Local authorities can leverage AI to analyse historical trends, monitor public sentiment, and identify gaps in service delivery. By combining data from sources like the Social Value Engine, councils can design inclusive, evidence-based policies that respond to community needs.

Leveraging the Social Value Engine's Curated Sources

The Social Value Engine provides public sector bodies with a robust platform for measuring social value, backed by around 700 curated data sources, or "proxies". This extensive collection ensures that social return on investment calculations can be both accurate and transparent, tracing each outcome to a specific, verified piece of research. Because these data sources have already been rigorously vetted for relevance and reliability, public sector organisations can confidently select

those best suited to assessing the true impact of their initiatives.

By eliminating the need to identify and verify data independently, the Social Value Engine frees up valuable time and resources for more meaningful work. Its sources align with internationally recognised social value frameworks, making it easier for councils to meet reporting and compliance obligations. Moreover, the platform's ongoing updates provide access to current data, ensuring that strategic decisions remain informed and evidence-based.

Using AI to Improve Stakeholder Data Collection and Analysis

A Theory of Change framework, along with early stakeholder identification, sets the foundation for effective community projects. By integrating AI into this process, public sector bodies can gain faster, more accurate insights into how their decisions affect local populations.

1. **Review Past Data**
   AI can quickly sift through historical project information, detecting trends that might otherwise go unnoticed. This

insight helps predict potential outcomes of future initiatives.

2. **Monitor Real-Time Sentiment**
   Tools that analyse social media posts or emails provide immediate feedback on public opinion, enabling timely responses to emerging concerns.

3. **Refine Surveys and Focus Groups**

   a. Adaptive Questions: AI can tailor survey questions in real time, ensuring they remain relevant to respondents.

   b. Thematic Analysis: By rapidly identifying main themes in focus group discussions, AI reduces the time required for manual analysis.

4. **Incorporate Novel Data Sources**
   Integrating information from sources like smart sensors offers hard evidence of changes in local environments, lending greater accuracy to assessments.

5. **Detect Gaps and Inequities**
   AI can flag whether certain groups are being overlooked during data collection,

promoting a more inclusive picture of community needs.

6. **Facilitate Continuous Learning**
   Complex data is translated into accessible lessons, helping teams refine their approaches and share knowledge more effectively.

7. **Forecast Potential Outcomes**
   AI can predict how stakeholders might respond to different strategies, guiding more targeted decision-making.

8. **Automate Reporting**
   Generating regular stakeholder impact updates saves considerable time and ensures key findings are communicated promptly.

9. **Uncover Hidden Connections**
   By examining large datasets, AI might reveal unexpected relationships between groups or outcomes, helping stakeholders better understand community dynamics.

10. **Optimise Communication**
    Using historical engagement data, AI can suggest the most effective ways to reach different stakeholder groups.
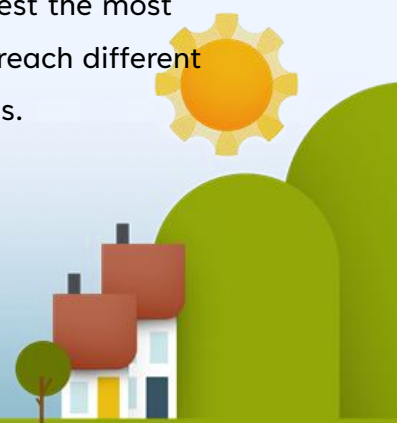
Chapter 5: A (Very Brief) Introduction to Data Ownership and Privacy

When adopting AI technologies—particularly in the public sector—it is vital to address data ownership and privacy to ensure that sensitive information remains secure and accessible only to those with legitimate authority. This chapter provides a concise overview of the main considerations in these areas, serving as an entry point to more detailed legal and technical discussions.

It is important to emphasise that this chapter offers only a starting point for tackling these complex issues. Legal and IT experts in public sector bodies must take the lead in forming robust policies, frameworks, and procedures that comply with current regulations and best practices.

In the sections that follow, we will examine the significance of data ownership, explore potential risks arising from data sharing, discuss privacy concerns, and highlight the need for transparency and human oversight in AI systems. These insights are meant to spark deeper conversations within your organisation and guide future policy development.

Data Ownership and the Risk of (Inadvertently) Sharing Private Data

Establishing clear policies on data ownership is essential for public sector bodies, as it defines who has rights to the collected and processed data, how that data may be used, and who can access it. A good starting point is to create a comprehensive data inventory, cataloguing each type of data used in AI systems, along with its source and a designated owner within the organisation. This inventory should be reviewed and updated regularly to maintain its relevance. Data ownership clauses also need to be addressed in contracts with AI service providers, specifying that the public sector body retains ownership of all data processed by the AI system—including any insights or outputs generated by the AI.

Sharing data with large technology companies or other external AI platforms poses inherent risks, as these entities may follow different data handling practices. Once information enters systems owned by third parties, the original controller often loses visibility and control, especially if the data is used to further train AI models. There is also the possibility of data breaches, despite robust security measures. To mitigate these risks, carefully vet potential AI service providers. Look for strong data protection policies, transparent handling practices, and a demonstrated track record of responsible data management. Using anonymised or synthetic data, where possible, further reduces exposure of sensitive information.

Public sector bodies must also address privacy concerns tied to AI's capacity for processing vast amounts of data and identifying patterns that can inadvertently reveal personal details. Adopting strong data governance practices—such as data minimisation and purpose limitation—helps safeguard residents' privacy. Regular privacy impact assessments are advisable when implementing new AI systems or significantly altering existing ones.

Organisations must also consider whether to host AI systems locally or use public cloud services. While cloud providers offer scalability and advanced security features, local hosting provides greater control over data and access. Ultimately, the decision depends on available resources, technical expertise, and the

sensitivity of the data involved. If opting for cloud services, selecting providers that comply with relevant regulations and offer robust data protection measures is crucial. In some cases, a hybrid approach—keeping the most sensitive data on local servers while using the cloud for less critical tasks—may strike the right balance between security and efficiency.

Maintaining Data Privacy

Data privacy should be a central consideration in all AI-related activities. While compliance with relevant regulations is essential, it can be bolstered by additional protective measures. Organisations can incorporate encryption, perform regular security audits, and implement strict access controls. Using data anonymisation techniques and establishing clear data retention and deletion policies further strengthens these safeguards. Regular reviews of data usage and access can help ensure that information is employed as intended and can also catch any unusual patterns early. Staff training on data protection principles ensures that everyone within the organisation

understands their responsibilities in maintaining data security.

The Importance of Transparency

When AI systems play a part in decision-making, there should be a way to explain how those decisions or recommendations are reached. Exploring explainable AI techniques can offer insights into these processes, which is especially valuable when outcomes significantly affect individuals or communities. By demystifying AI operations and clarifying how outputs are generated, public sector bodies can help build trust and confidence in AI-driven initiatives.

Maintaining Human Oversight

Encouraging staff to question AI insights helps identify errors or biases. Regularly updating policies and staying informed on advancements will help local authorities maximise AI's benefits while protecting privacy and community interests.
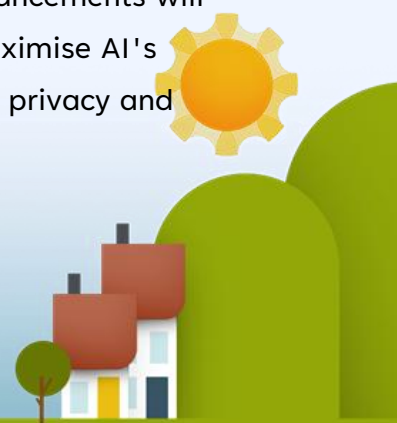
Conclusion: Building a Better Future with Ethnical AI

AI is not just a tool for efficiency; it is a catalyst for reimagining how local authorities address societal challenges and enhance community well-being. The key lies in how it is deployed—not just to automate processes but to amplify human strengths, foster inclusion, and deliver meaningful, long-term value.

As AI becomes embedded in public sector strategies, its success will hinge on a commitment to fairness, transparency, and collaboration. By integrating ethical principles into its design and use, AI can become a bridge between data-driven insights and the lived experiences of diverse communities. This requires local authorities to remain adaptable, continually questioning the assumptions behind AI-driven decisions and ensuring these technologies work for everyone.

AI's promise is not in replacing human judgement, but in elevating it—helping local governments make more informed, empathetic, and impactful choices. By anchoring AI initiatives in the principles of social value and equity, we have the opportunity to shape a future where technology serves as a true partner in creating resilient, connected, and thriving communities.

Appendix A: Generative AI Tools

There are countless generative AI tools available to the public, many of which use the same large language models under the hood. Below is a brief overview of some well-known tools and platforms, grouped by their primary focus.

Conversational

OpenAI GPT-4 / ChatGPT

An advanced language model capable of generating human-like text for various tasks, including content creation, question answering, and conversational interactions.

- **Website:** openai.com
- **Owner:** OpenAI
- **Pricing:** Subscription-based (ChatGPT offers a free tier with limited capabilities and a paid tier for full access)

Gemini

A generative AI model developed by Google, aimed at creating advanced text outputs for applications ranging from content creation to conversational agents.

- **Website:** gemini.google.com
- **Owner:** Google
- **Pricing:** Free

Copilot

An AI-powered tool that integrates with the Bing search engine, offering a new type of search experience through code assistance and other advanced features.

- **Website:** copilot.microsoft.com
- **Owner:** Microsoft
- **Pricing:** Free (paid Pro version available)

Claude.ai

A conversational AI model designed to generate human-like text and offer assistance in a chat-style environment.

- **Website:** claude.ai
- **Owner:** Anthropic
- **Pricing:** Free (paid Pro version available)

Images

DALL-E 2

An AI model that generates images from textual descriptions, enabling the creation of unique artwork based on user prompts.

- **Website:** openai.com/dall-e-2
- **Owner:** OpenAI

- **Pricing:** Subscription-based (free credits available for limited use)

DeepArt

Transforms photos into artwork inspired by famous artists' styles.

- **Website:** deepart.io
- **Owner:** DeepArt
- **Pricing:** Subscription-based

NightCafe Creator

Creates artwork from textual descriptions using neural style transfer and other algorithms.

- **Website:** creator.nightcafe.studio
- **Owner:** NightCafe Studio
- **Pricing:** Subscription-based (free tier available)

MidJourney

Generates high-quality images from textual prompts, often used for concept art and visual storytelling.

- **Website:** midjourney.com
- **Owner:** MidJourney
- **Pricing:** Subscription-based

Stable Diffusion

A deep learning, text-to-image model for generating detailed images from text descriptions, also applicable to tasks like inpainting and outpainting.

30

- **Website:** stability.ai
- **Owner:** Stability AI
- **Pricing:** Free (open-source model)

Music

AIVA

An AI composer that creates music for various purposes, from video game soundtracks to advertisements.

- **Website:** aiva.ai
- **Owner:** AIVA Technologies
- **Pricing:** Subscription-based (limited free tier)

Amper Music

An AI-driven music composition tool for creating original music tracks quickly and easily.

- **Website:** ampermusic.com
- **Owner:** Amper Music
- **Pricing:** Subscription-based

Suno

An AI music generation platform for creating and customising music tracks using advanced algorithms.

- **Website:** suno.ai
- **Owner:** Suno AI
- **Pricing:** Subscription-based (limited free credits)

Udio

Generates and customises music based on user preferences and requirements.

- **Website:** udio.ai
- **Owner:** Udio AI
- **Pricing:** Subscription-based

Video

Synthesia

An AI platform that generates videos with lifelike avatars from text input, streamlining content creation.

- **Website:** synthesia.io
- **Owner:** Synthesia
- **Pricing:** Subscription-based

Pictory

Creates and edits videos from text, offering automated features like voiceovers and subtitle generation.

- **Website:** pictory.ai
- **Owner:** Pictory
- **Pricing:** Subscription-based

Lumen5

Transforms blog posts and articles into engaging videos powered by AI.

- **Website:** lumen5.com
- **Owner:** Lumen5

- **Pricing:** Subscription-based (free tier available)

3D Models

Runway ML

A platform for creating and editing 3D models, animations, and other visual elements using machine learning tools.

- **Website:** runwayml.com
- **Owner:** Runway
- **Pricing:** Subscription-based (free tier available)

GANPaint Studio

Allows users to modify and create 3D models and scenes via generative adversarial networks (GANs).

- **Website:** ganpaint.io
- **Owner:** MIT-IBM Watson AI Lab
- **Pricing:** Free

ArtBreeder

A collaborative AI platform for creating and evolving images, including 3D models.

- **Website:** artbreeder.com
- **Owner:** ArtBreeder
- **Pricing:** Subscription-based (free tier available)

Appendix B: Glossary of Terms

Artificial Intelligence (AI)

A branch of computer science focused on creating systems capable of performing tasks that typically require human intelligence, such as understanding natural language, recognising patterns, and making decisions.

Generative AI

A subset of AI that uses algorithms and models to create new content, such as text, images, music, or code, often mimicking human creativity.

Large Language Model (LLM)

A type of generative AI model, typically based on deep learning, trained on vast amounts of text data to understand, generate, and manipulate human language.

Neural Network

A computational model inspired by the human brain, consisting of interconnected nodes (neurons) that process data and learn patterns.

Deep Learning

A subset of machine learning involving neural networks with many layers, enabling complex data representations and high-level abstractions in tasks like image and speech recognition.

Training Data

The dataset used to train an AI model, composed of numerous examples that help the model learn patterns and relationships in the data.

Natural Language Processing (NLP)

A field of AI concerned with the interaction between computers and humans through natural language, enabling machines to understand, interpret, and generate human language.

GPT (Generative Pre-trained Transformer)

A series of large language models developed by OpenAI, known for their ability to produce coherent and contextually relevant text.

BERT (Bidirectional Encoder Representations from Transformers)

An LLM designed by Google, using the transformer architecture to better understand the context of words in search queries and improve search result accuracy.

Prompt

A piece of text input given to a generative AI model to guide its text generation process, helping to produce relevant, context-specific output.

Self-Attention

A mechanism within transformer models that allows them to weigh the importance of different words in a sentence, capturing context and dependencies more effectively.

Context Window

The range of text an AI model considers when generating responses, influencing how it maintains coherence and relevance over longer passages.

Bias in AI

The presence of systematic, unfair discrimination in AI models—often arising from biased training data—which can lead to skewed or prejudiced outcomes.

Ethics in AI

The study and implementation of moral principles in AI development and deployment, ensuring responsible, fair, and harm-free use of technology.

Zero-Shot Learning

The ability of some AI models to perform tasks they were not explicitly trained for, by leveraging generalized knowledge from related tasks.

Multimodal AI

AI systems that process and generate data across multiple formats—such as text, images, and audio—providing richer, more comprehensive outputs.

Reinforcement Learning

A machine learning technique where an agent learns to make decisions based on rewards or penalties from its actions.