

Policy briefing

Digital ethics

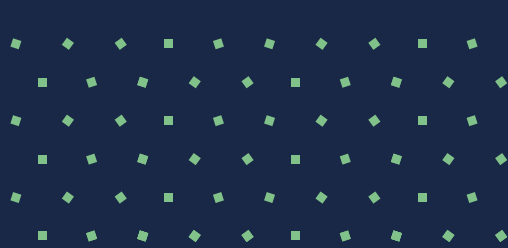
The ethical use of emerging technologies and data

September 2020



Table of contents

Introduction	04
<hr/>	
Understanding digital ethics	06
<hr/>	
Defining digital ethics	06
Digital ethics in context	07
Ethics by design	07
Ethics of use	08
Societal ethics	09
<hr/>	
Ethical principles and core values	10
<hr/>	
Emerging principles	10
Ethical framework for Smart Information Systems	11
Ethics and data protection in artificial intelligence	11
Place-based ethical challenge	11
Thinking globally – acting locally	12
Towards common core values	14
<hr/>	
Attributes of ethical practice	14
<hr/>	
Attribute 1: beneficence = do good	14
Attribute 2: non-maleficence = do no harm	15
Attribute 3: autonomy = preserve human agency	16
Attribute 4: justice = be fair	17
Attribute 5: explicability = operate transparently	18
<hr/>	



Conclusion	20
References	21
Appendix 1: emerging ethical standards	21
Appendix 2: emerging guidance	24

Introduction

The pace of local adoption and development of new technology solutions and data techniques (so-called Smart Information Systems) to “fix the plumbing” in the face of Covid-19 has been frenetic. Time and again during the crisis local communities have been called upon to fill the breach. Be it developing workable local test and trace systems or undertaking data analysis to fill the gaps in knowledge about specific localities, local public service leaders have been asked to re-think the “art of the possible” in relation to whether and how they should use technology and data.

Nevertheless, as headlines featuring algorithmic bias and discriminatory facial recognition show, the deployment of such solutions and approaches is not without risk. From time to time, they have opened up organisations to serious charges of bias, discrimination and injustice in the development, design and delivery of public services.

Such challenges highlight that re-thinking the art of the possible in this way is never easy. But, as Covid-19 has shown, they are part and parcel of the new normal for local public service leaders, policy makers and practitioners, in partnership with their wider communities’ needs, to meet the future head-on.

They also serve as an important reminder that a crisis is not a time to throw ethics out of the window. Rather, it is a time to take into consideration official guidance¹ to the public sector that stresses how the use of AI and data analytics should be proportionate and ethical, rather than simply the automatic first choice in all circumstances and situations.

This means it is time to look afresh and ensure that ethical considerations are embedded in how organisations develop and use technology. In order to do this successfully it is essential that we learn to identify, define and understand not only the potential benefits

of AI and data analytics but also the technological and algorithmic risks from the outset, so that we avoid discriminatory and adverse personal outcomes.

We call for more than just a tick-list approach. Local public services need to understand the wider ethical landscape and champion digital ethical practice at the very heart of place-based approaches to the responsible use of technology and data for the public good.

Above all, it requires us to understand and promote the ethical use of emerging technologies; proactively looking afresh at how we use the data they generate and store, and the public service designs, processes and interactions they enable; and considering how all this applies to the outcomes they generate, and how we can ensure public benefit by successfully addressing societal challenges, supporting planetary sustainability and minimising unintended consequences, as advocated in Kate Raworth’s concept of [Doughnut Economics](#).

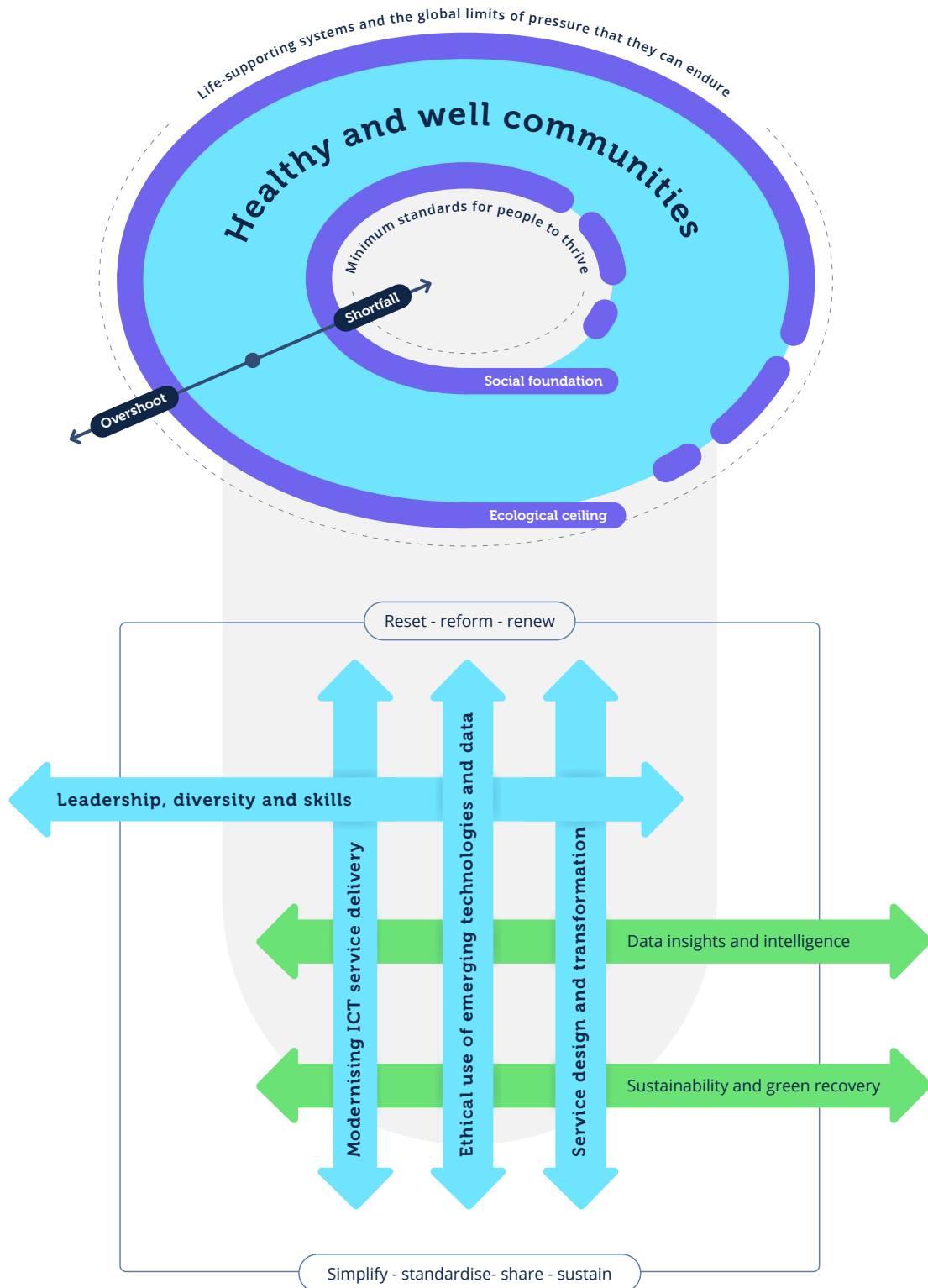
Socitm’s five key policy themes, including ethical use of emerging technologies and data – digital ethics – are presented by reference to the doughnut framework in Figure 1 on the following page.

Purpose and audience for this policy briefing

Socitm produces policy briefings on behalf of those working in the public sector, in particular local government. The focus of this briefing will be of interest to senior policy-makers, decision-makers and managers involved in the delivery of local public services.

Figure 1.

Socitm's ethical, digital, place-making model



Understanding digital ethics

Defining digital ethics

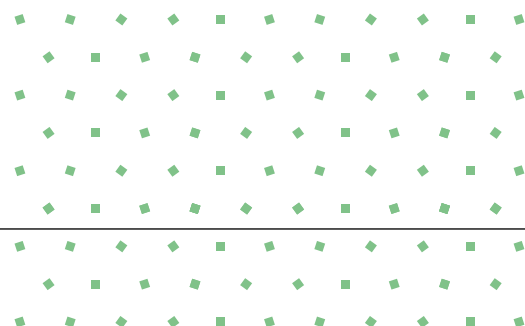
Taking a step back can help us focus on the real role of ethics. Quite simply, these are moral principles that govern a person's or group's behaviour and inform decision making. The basic role of ethics can be described as:

- › Establishing a foundation for decision making based on a core set of values committed to “doing no harm” and supporting human flourishing and environmental sustainability.
- › Providing a structure to address value dilemmas and conflicts that are specific to each context, on a case-by-case basis.
- › Drawing on the ideas built around values, deliberations and recommendations that support the common good.
- › Setting a baseline for choosing how best to consider an issue and then act on a case-by-case basis by rationally considering different and conflicting values/principles.

At a basic level it is possible to see digital ethics as “principles” and “concepts” that can be used to govern technology and data including factors like risk management and individual rights. In effect, they can be used to understand and resolve moral issues that have to do with the development and application of differing technology and data solutions and approaches across a range of ethical challenges such as: ²

- › **Access rights:** access to empowering technology as a right
- › **Accountability:** decisions made for who is responsible when considering success or harm in technological advancements

- › **Digital rights:** protecting intellectual property rights and privacy rights
- › **Environment:** how to produce technology that could harm the environment
- › **Existential risk:** technologies that represent a threat to the global quality of life pertaining to extinction
- › **Freedom:** technology that is used to control a society raising questions related to freedom and independence
- › **Health and safety:** health and safety risks that are increased and imposed by technologies
- › **Human enhancement:** human genetic engineering and human-machine integration
- › **Human judgement:** when can decisions be judged by automation and when do they acquire a reasonable human?
- › **Over-automation:** when does automation decrease quality of life and start affecting society?
- › **Precaution principle:** who decides that developing this new technology is safe for the world?
- › **Privacy:** protection of privacy rights
- › **Security:** is due diligence required to ensure information security?
- › **Self-replicating technology:** should self-replicating be the norm?
- › **Technology transparency:** clearly explaining how a technology works and what its intentions are
- › **Terms of service:** ethics related to legal agreements



In effect the ethics of digital technology and data can be said to focus on the ethical aspects of technological design and use, together with the ethical impacts of digital technology on society as whole, as follows: ³

- › **Ethics by design:** these focus on the design phase of digital and data tools.
- › **Ethics of use:** these aim to examine how the service users and employees as well as the managers and partners of an organisation use emerging technology and data.
- › **Societal ethics:** these examine the impacts of digital technology and data analytics on wider society.

Digital ethics in context

These three categories of ethical focus are of course interconnected, but we need to understand context in which they operate in order to better define the basic issues at stake, and how they can start to be addressed. Taking each of these three categories in turn and looking at the following set of introductory questions and points to consider can help to build up a clearer understanding of the digital ethical landscape and how organisations and individuals can start to navigate it as they consider these and other points as they emerge in their wider examination of these and related issues.

The following **points to consider** are by no means exhaustive so should be seen as a starting point to further discovery...

Ethics by design

These focus on the design phase of digital and data tools. It directly concerns technology in all its technical complexity and the know-how of engineers, programmers, etc. These ethics therefore touch in particular on the deontology (duty-based ethics) of

digital creators of all kinds (developers, digital designers, project managers, etc.). Indeed, they have an ethical responsibility from the design stage onwards, insofar as data or algorithms may or may not reproduce human biases, reveal new discriminations (or reproduce them on a larger scale), give rise to injustices, etc.

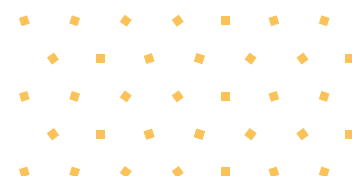
Are the solutions and approach under examination proportionate and ethical?

- › Ensure the system is appropriate for the application (the function, the problem) under consideration
- › Will the solution work and be better than anything else for that purpose?
- › Take care to make sure its use in the proposed context be lawful, safe, acceptable to stakeholders, wise, and not have bad side effects

Do you understand the data context and are you looking to rigorously test it is critical to the integrity of any process?

Data use needs to be examined on a case by case basis by addressing the following points in tandem with the proportionality question above.

- › Are the precise calculations and data processing methods used in the program?
- › What is the exact nature of the data used in the development and testing of the program, and in its intended operation?
- › What is the purpose for which the system will be used?
- › Remember that the social, political, professional, environmental and operational process and practice contexts within which it will be used – its use will change these, and these are where consequences will be felt.



Are the risks of bias in the datasets identified and addressed?

- › Train policy makers, designers and practitioners to focus on identifying, defining and understanding the potential of algorithmic risks from the outset if we are serious about avoiding discriminatory and adverse personal outcomes.
- › Carry out rigorous impact assessment (in accordance with official guidance and GDPR) in order to analyse the possible design-induced discriminatory impacts of algorithms
- › Put in place checks and balances at each stage of development to ensure there is no bias in the results

Can the operating rationale of the algorithms deployed for artificial intelligence be explained?

- › Have a system explain ability policy encompassing the whole chain (data provenance, explanation of the reasoning followed)
- › Develop algorithms that are transparent by design, in order to make it easier to explain them and to analyse how they reason
- › Adopt a labelling (with an ethical scoring/rating system) and ethical support approach

Does the organisation offer training programmes on ethics in the creation of digital tools?

- › Set up training workshops and/or skills refresher courses within the IT and related service departments

Are solutions designer's representative of the social, ethnic and gender diversity of society?

- › Draw up a HR policy ensuring social and gender diversity in the workplace

Are new projects evaluated for their impact on privacy and personal data?

- › Consider setting up an ethics committee to approve sensitive projects

Do tools and solutions protect personal data by design and factor in the right to be forgotten into the design chain?

- › Adopt a privacy by design approach, in accordance with the requirements of the GDPR: this means building the protection of personal data into products and services by design, but also by default (notably by abiding by the data minimisation principle introduced by the GDPR).
- › Remember these are also cultural challenges because this concept needs to be factored into a project early on

Does the correlation of data collected from various sources result in the production of personal information (as part of big data and AI projects, for instance)?

- › Put in place a system that measures the personalisation of data after processing operations

Ethics of use

These aim to examine how the service users and employees as well as the managers and partners of an organisation use emerging technology and data. This entails conducting an ethical evaluation of how people use the technological resources at their disposal.

Is there a robust set of checks and balances built around political and executive scrutiny?

- › Focus on the adoption of due diligence frameworks, appropriate standards, principles together with accountable public good focussed audit and risk regimes that allow for an effective measure of public participation in all the stages from initial evaluation to implementation

Are ethical rules for data collection and processing defined and shared internally within the organisation?

- › Raise staff awareness with data ethics training and workshops

Is there a framework for internal rights of access to personal and/or sensitive data?

- › Clearly define procedures for access to sensitive data based on employees' profiles and roles

Are digital tools designed with the accessibility needs of disabled people in mind?

- › By default, design solutions that are accessible for people with disabilities

Are ethics-related issues addressed on a cross-functional basis within the organisation?

- › Consider establishing a Chief Digital Ethics Officer tasked with ensuring the overall coherence of the organisation's "ethics and digital" policy
- › Put in place an awareness-raising programme for all employees (information and examples of best practice)

Are employees informed of how their data will be stored and processed, and their rights in this area?

- › Inform employees of how their data will be stored and processed, and their rights in this area (display, updating of internal regulations)

Are the consequences of the internal use of certain digital tools assessed?

- › Carry out an assessment of the impact of digital tools on the day-to-day experience of employees in the organisation.

Are the users of personalised services given the option to manage their settings?

- › Ensure that the information given to users is clear and transparent
- › Make it easy for users to change their personal data management settings, and to make informed choices

Are users informed of the terms of use of a digital solution or application?

- › Consider drawing up a digital user charter setting out the ethical terms of use of a solution
- › Look at setting out a framework for the use of a solution in contracts, and allow for stakeholder object to noncompliant use

Is there a policy allowing you to check, when digital solutions are built by a number of partners, that the whole process is ethical?

- › Ensure that the ecosystem is trustworthy and give each partner a vision of the purpose of the overall solution
- › Call on trusted third parties, certifications and/or labels, demonstrating the ethical commitment of each participating stakeholder

Societal ethics

These examine the impacts of digital technology and data analytics on wider society. It thus deals with the acceptability of digital innovations and solutions, human rights and agency, the environmental/energy footprints of digital tools, and wider issue of social inclusion.

Has an approach designed to improve the environmental footprint of the IT system been introduced?

- › Identify a manager and draw up an action plan including raising the awareness of all IT department staff and users, building on recognised standards
- › Factor in environmental impact when entering into any contract that has consequences for the environmental footprint of the IT system

Does the evaluation of the IT system's environmental impacts cover primary energy, greenhouse gas emissions, water, the depletion of abiotic resources, paper and WEEE (waste electrical and electronic equipment)?

- › Conduct a regular assessment of the environmental footprint of the IT system based on recognised and auditable indicators

Is the societal impact of projects (origin of materials, partners' good practices) taken into account?

- › Carry out a societal impact assessment of projects

Is an assessment carried out of the impact of innovations on jobs in the company, especially when automation occurs?

- › Forecast, with the help of teams specialising in forward planning and strategy, the impacts of technological change on the organisation's jobs and activities.
- › Include the impacts of automation and more broadly of digital technology in strategic workforce planning.

Are addiction phenomena factored into the design of digital solutions?

- › Formally discourage the use of "dark patterns" (interface design tricks intended to trick the user)

Are the risks of human cognitive biases factored into the design of digital solutions?

- › Ensure that digital applications and solutions have not been designed in such a way as to deliberately manipulate users by exploiting cognitive biases

Ethical principles and core values

Emerging principles

Given the global nature of this challenge, the [Organisation for Economic Co-operation and Development](#) (OECD) have been leading on the development of a core set of digitally ethical principles for wider international adoption which comprise the following themes (see also [OECD recommendations on AI](#)⁴):

- › **Inclusive growth, sustainable development and well-being:** this highlights the potential for trustworthy AI to contribute to overall growth and prosperity for all – individuals, society, and planet – and advance global development objectives.
- › **Human-centred values and fairness:** AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and should include appropriate safeguards to ensure a fair and just society.
- › **Transparency and explainability:** this principle covers 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 and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the other values-based principles for AI.



These principles have been adopted by the UK and all the other [G20](#) countries and are now being taken up by a growing number of international bodies, UN member states (via the [UNESCO AI ethics consultation](#)), informing the World Economic Forum [AI and Robotics forward programme](#). They also form the basis of the new [Global Partnership on Artificial Intelligence](#) (GPAI) which the UK, US, EU and other partners have established to champion responsible AI and data governance.

Ethical framework for Smart Information Systems

At an operational level, the [European High-Level Expert Group](#)⁵ and [Sherpa programme](#)⁶ have likewise identified the following key requirements for assessing ethical practice across smart information systems:

- **Human agency and oversight:** including fundamental rights, human agency and human oversight
- **Technical robustness and safety:** including resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility
- **Privacy and data governance:** including respect for privacy, quality and integrity of data, and access to data
- **Transparency** Including traceability, explainability and communication
- **Diversity, non-discrimination and fairness:** including the avoidance of unfair bias, accessibility and universal design, and stakeholder participation
- **Societal and environmental wellbeing:** including sustainability and environmental friendliness, social impact, society and democracy
- **Accountability:** including auditability, minimisation and reporting of negative impact, trade-offs and redress

Ethics and data protection in artificial intelligence

Alongside this the [International Conference of Data Protection and Privacy Commissioners](#)⁷ (which includes the UK ICO) has initiated a declaration on ethics and data protection in artificial intelligence that outlines the following guiding principles as its core values, to preserve human rights in the development of AI and Data analytics:

- AI and machine learning technologies should be designed, developed and used in respect of fundamental human rights and in accordance with the fairness principle.
- Continued attention and vigilance, as well as accountability, for the potential effects and consequences of, artificial intelligence systems should be ensured.
- AI systems transparency and intelligibility should be improved, with the objective of effective implementation.
- As part of an overall “ethics by design” approach, artificial intelligence systems should be designed and developed responsibly, by applying the principles of privacy by default and privacy by design.
- Empowerment of every individual should be promoted, and the exercise of individuals’ rights should be encouraged, as well as the creation of opportunities for public engagement.
- Unlawful biases or discriminations that may result from the use of data in artificial intelligence should be reduced and mitigated.

Place-based ethical challenge

As Socitm’s work on digital ethics demonstrates, organisations need to identify what is unique and different about an ethical, place-based approach to using emerging technologies and data. Whilst

at the same time explore how adopting a digitally ethical approach can help to design better services, leveraging benefits and better outcomes.

The Socitm suite of 'smart places' and location intelligence guides, neatly demonstrate how emerging technology and data are connecting place-based infrastructure, business, communities, public service and individual citizens in ways that were previously impossible, such as:

- › **Jobs:** improved access for people in areas with persistent inter-generational unemployment.
- › **Housing:** accessible and varied.
- › **Education:** availability of high-quality education and sustained attainment.
- › **Public health:** healthy and safe environments and behaviours.
- › **Social care:** improved coordination with health and police, earlier and better interventions for troubled families and the elderly.
- › **Fraud reduction:** identifying potential council tax and housing benefit fraud by finding signature "spatial patterns" in big data sets.
- › **Businesses:** sustained conditions for businesses and communities to innovate and thrive.
- › **Smarter planning:** smarter design and construction of new developments and better integrated transport for sustainable living.
- › **Arts and culture:** valuing diversity and facilitating access for all.
- › **Changing community engagement:** through social media and smartphone apps that give people a more attractive and interactive experience, including 3D visualisation.

And as Socitm's *Planting the flag*⁸ - a new local normal illustrates, drawing ideas and experience from members of Socitm and its partner associations (LOLA and MCE)

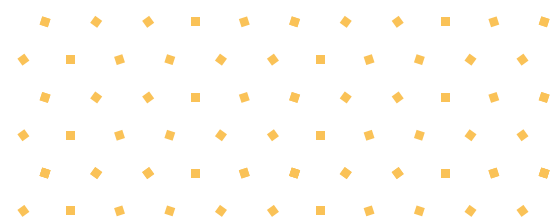
around the world. The following eight key areas of place-based activity are being transformed by local responses to Covid-19 and potentially offer a model of place based ethical change built around key "principles/ concepts" of Simplify – Standardise – Share – Sustain:

- › Place-based leadership
- › Democratic engagement and renewal
- › Service design
- › Workstyles
- › Virtual infrastructure
- › Data insights
- › Asset rationalisation
- › Living spaces.

Thinking globally – acting locally

More widely emerging thinking around what constitutes a place-based approach to the ethical use of emerging technology and data analytics is being stimulated by initiatives like the UN's sustainable development goals, the UK's Local Digital Declaration and the World Economic Forum's work on economic recovery from Covid-19.

Applying these lessons from this thought leadership and Socitm's to digital placemaking can if handled aright, boost the social, cultural, environmental and economic value of places by using location-specific digital technology to foster deeper relationships between people and the places they inhabit. Such an approach is at the heart of the emerging concepts



of [Doughnut Economics](#) and [City Portraits](#), which seeks to model ethical ways in which “people and planet can thrive in balance” and builds on the earlier initiatives such as International City/County Management Association (ICMA) “[Building Digitally Inclusive Communities](#)” principles below:

- › **Availability and affordability:** communities need reliable and affordable access to broadband technology infrastructure in order to be fully-engaged and competitive in today’s information-based world.
- › **Public access:** in a world connected by technology, all people, regardless of income, need access to information and communication technologies in order to be fully engaged members of society, both economically and socially.
- › **Accessibility for people with disabilities:** communities should ensure the full participation of all their members, by embedding accessibility to digital technology for people with disabilities throughout their institutions, processes, and public awareness efforts.
- › **Adoption and digital literacy:** beyond having access to technologies, people, businesses, and institutions need to understand digital technologies and how to use them effectively to achieve their educational, economic, and social goals.
- › **Consumer education and protection:** consumers – both individual and institutional – need accurate, unbiased information to understand the technology options available to them, including how to buy and maintain equipment and how to safely navigate the digital world.
- › **Education:** educational institutions should ensure that students have the digital skills to fill the jobs of today and tomorrow, and to reap the potential rewards of lifelong digital learning.
- › **Economic and workforce development:** technology is a powerful engine of innovation and economic growth in today’s world. For individuals and businesses to succeed in this

environment, communities need to foster the mastery of 21st century skills and encourage use of technology for economic development.

- › **Civic engagement:** residents should be easily able to interact electronically with institutions, government, and one another, to participate actively in community affairs.
- › **Public safety and emergency services:** communities can increase their emergency responsiveness through effective deployment of digital technologies, ensuring the public the best possible emergency preparedness
- › **Health care:** communities should have the digital technologies necessary to support the health care needs of their populations, especially in areas with limited health care facilities, to afford all their members access to the best possible health care.
- › **Quality of life:** individual members of a community should have access to technologies that promote social engagement and the pursuit of productive and creative interests.

This requires places to start to identify what is unique and different about an ethical, locally-based approach to using smart information systems, in order to help us design better services, leveraging benefits and better outcomes. This approach is typified by Amsterdam’s environmental and societal wellbeing City Doughnut initiative⁹ (that in partnership with the Doughnut Economics Action Lab) is applying the following principles for putting ethical place-based change into practice:

- › **Embrace the 21st century goal.** Aim to meet the needs of all people within the means of the living planet. Seek to align your organisation’s purpose, networks, governance, ownership and finance with this goal. Expect the work to be challenging, innovative and transformative.
- › **See the big picture.** Recognise the potential roles of the household, the commons, the market and the state - and their many synergies - in transforming economies. Ensure that finance serves the work rather than drives it.

- **Nurture human nature.** Promote diversity, participation, collaboration and reciprocity. Strengthen community networks and work with a spirit of high trust. Care for the wellbeing of the team.
- **Think in systems.** Experiment, learn, adapt, evolve, and aim for continuous improvement, Be alert to dynamic effects, feedback loops and tipping points.
- **Be distributive.** Work in the spirit of open design and share the value created with all who co-create it. Be aware of power and seek to redistribute it to improve equity amongst stakeholders.
- **Be regenerative.** Aim to work with and within the cycles of the living world. Be a sharer, repairer, regenerator, steward. Reduce travel, minimize flights, be climate and energy smart.
- **Aim to thrive rather than to grow.** Don't let growth become a goal in itself. Know when to let the work spread out via others rather than scale up in size.

Towards common core values

The UK is playing a leading role in these agendas and official bodies like the Office for Artificial Intelligence, Centre for Digital Ethics and Information Commissioner's Office are working closely with Digital Ethics Lab, Alan Turing Institute, Open Data Institute and Digital Catapult in championing digitally ethical practice across the UK public sector. As a result, we are seeing an emerging set of common core values¹⁰ or attributes that can help inform digital ethical practice:

- **Beneficence: do good.** Benefits of work should outweigh potential risks.
- **Non-maleficence: do no harm.** Risks and harms need to be considered holistically, rather than just for the individual or organisation.

- **Autonomy: preserve human agency.** To make choices, people need to have sufficient knowledge and understanding.
- **Justice: be fair.** Specific issues include algorithmic bias and equitable treatment.
- **Explicability:** operate transparently so as to explain systems working and its outputs.

Attributes of ethical practice

Understanding how these core values or attributes can help to make up a digitally ethical approach is therefore key to the work of informing and shaping organisations' strategy, culture and operational implementation in support of the ethical use of emerging technology and data.

Each attribute focuses on a number of themes and their relevance to a given outcome can be tested by applying the following set of scrutiny questions to see the extent that ethical practice has become successfully embedded in how organisations develop and use technology and become part of their operational, intellectual and systematic DNA.

In this way, taking the five core values or attributes of Beneficence, Non-Maleficence, Autonomy, Justice and Explicability in turn, it is possible to scope out the basis of digital ethical practice framework for working with Smart Information Systems as follows:

Attribute 1: beneficence = do good

Benefits of work should outweigh potential risks. Digital, data and technology that interact with people's wellbeing, finances, relationships and health particularly require robust ethical principles.

Key themes: human flourishing, well-being, dignity, common good, and sustainability.

- › Individual, societal, and environmental wellbeing: Sustainable and environmentally friendly AI and big data systems, individual wellbeing, social relationships and social cohesion, and democracy and strong institutions
- › Because AI and big data systems can have huge effects for individuals, society, and the environment, systems should be trialed, tested, and anomaly-detected to ensure the reduction, elimination, and reversal of harm caused to individual, societal and environmental well-being

Areas of focus

- › **Requirements stakeholder participation:** to develop systems that are trustworthy and support human flourishing, those who will be affected by the system should be consulted
- › **Protection of fundamental rights** so that the exercise of individuals' rights should be encouraged, as well as the creation of opportunities for public engagement.
- › **Sustainable and environmentally friendly:** the system's supply chain should be assessed for resource usage and energy consumption
- › **Justification:** the purpose for building the system must be clear and linked to a clear benefit – systems should not be built just for the sake of it.

Beneficence scrutiny questions

How do we keep “do good” ethics in the spotlight?

- › **Remember:** ethics is a pervasive aspect of good practice; ethical issues are intrinsically involved in ensuring good use of data and technologies.
- › **This means:** ethical considerations are ubiquitous and intrinsic to design, development, deployment and delivery, and individuals and organisations need to keep the ethics in the spotlight at all time.

How do we prioritise human lives and interests in adopting emerging solutions?

- › **Remember:** Digital, Data and Technology (DDaT) interacts with the most important aspects of people's lives: their wellbeing, their finances, their social relationships, and their emotional and mental health.
- › **This means:** Emerging approaches and solutions need to focus on people's needs, money and wellbeing with due care and attention; ensuring that robust ethical principles and standards are applied when developing technologies that touch these and other important aspects of people's lives.

Attribute 2: non-maleficence = do no harm

Risks and harms need to be considered holistically, rather than just for the individual or organisation. Privacy and security risks require social and organisational design, not just technological.

Key themes: safety, reliability, robustness, data provenance, privacy, and cybersecurity.

- › Technical robustness, safety in support of privacy and data governance: Including resilience to attack and security, fall-back plan and general safety, accuracy, reliability and reproducibility. Together with respect for privacy, quality and integrity of data, access to data, data rights and ownership.
- › Because we value humans, human life, and human resources, it is important that the system and its use is safe (often defined as an absence of risk) and secure (often defined as a protection against harm, i.e., something which achieves safety). Under this attribute we should also include the quality of system decisions in terms of their accuracy, reliability, and precision.
- › In particular, as systems use data that is private or sensitive, it is important to make sure that the system does not violate or infringe upon the right to privacy, and that private and sensitive data (such

as linked to an individual's ability to have a private life), is well-protected. This requires due diligence over the quality and integrity of data (i.e., whether the data is representative of reality), access to data, and the wider set of data rights such as ownership.

Areas of focus:

- › **Resilience to attack and security:** systems should be protected against vulnerabilities that can allow them to be exploited by adversaries.
- › **Fallback plan and general safety:** systems should have safeguards that enable a fallback plan in case of problems.
- › **Accuracy:** for example, documentation that demonstrates evaluation of whether the system is properly classifying results.
- › **Privacy and data protection:** systems should guarantee privacy and data protection throughout a system's entire lifecycle.
- › **Reliability and reproducibility:** does the system work the same way in a variety of different scenarios?
- › **Quality and integrity of the data:** when data is gathered it may contain socially constructed biases, inaccuracies, errors and mistakes – this needs to be addressed
- › **Social impact:** the effects of systems on people's physical and mental wellbeing should be carefully considered and monitored.

Non-maleficence scrutiny questions

How do we consider ethical risks and harms?

- › **Remember:** that involvement with the ethical issues concerning emerging technologies and big data don't go away once we have undertaken our own particular tasks and fulfilled our immediate responsibilities in the delivery chain.

- › **This means:** we need to view ethical risks and harms holistically – taking a comprehensive end-to-end view of the ethical issues across all stages of design, development, deployment and delivery and how they interact with people and society in an ethical manner.

How do we ensure design processes support privacy and security?

- › **Remember:** ethical design is not only technical design (of networks, databases, devices, platforms, websites, tools, or apps), but also social and organisational design of groups, policies, procedures, incentives, resource allocations, and techniques that promote privacy and security objectives.
- › **This means:** whilst DDaT implementation approaches and solutions will vary depending on context, it is an ethical imperative that the values of privacy and security are always at the forefront of operational design, planning, execution, and oversight.

Attribute 3: autonomy = preserve human agency

To make choices, people need to have sufficient knowledge and understanding. It is important to involve stakeholders and interest groups in ethical risk assessment and design.

Key themes: consent, choice, enhancing human agency and self-determination

- › Because we value the ability for humans to be autonomous and self-governing (positive liberty), humans' freedom from external restrictions (negative liberties, such as freedom of movement or freedom of association).
- › Underpinning this is the fact that each individual has an inherent worth and we should not undermine respect for human life (human dignity), we need to ensure that AI and big data systems do not negatively affect human agency, liberty, and dignity.

Areas of focus:

- › **Human agency users** should be able to make informed autonomous decisions regarding smart information systems
- › **Human oversight:** may be achieved through governance mechanisms such as human-on-the-loop, human-in-the-loop, human-in command
- › Ensure the **protection of the stakeholders' human agency** and positive liberty by keeping them informed, ensuring that they are neither deceived nor manipulated, and can meaningfully control the system;
- › Ensure the **protection of the stakeholders' negative liberty** by ensuring that they have the freedom to use the system and that they are not restrained in functionality and opportunity;
- › Ensure the **protection of the stakeholders' human dignity** by ensuring that the system is not used to directly or indirectly affect or reduce their autonomy or freedom, and does not violate their self-respect.

Autonomy scrutiny questions**How do we ensure that relevant stakeholders and interest groups in the design and decision-making are consulted?**

- › **Remember:** stakeholder and interest group involvement in ethical risk assessment and design is key to maintaining public and community confidence and trust in DDaT outcomes.
- › **This means:** it is important that stakeholder input does not simply reflect the same ethical perspectives as those already held within your organisation. External input from a more truly representative body of those likely to be impacted by specific DDaT outcomes is required.

How do we balance user expectations and the reality of what DDaT can deliver?

- › **Remember:** in creating DDaT approaches, consider how stakeholders' expectations of a particular solution may diverge from the reality of what can be delivered.
- › **This means:** we always have an ethical duty of care to ensure stakeholders are properly informed about not just the benefits but also the limitations and risks of a particular emerging approach or solution.

Attribute 4: justice = be fair

Specific issues include algorithmic bias and equitable treatment. Consider whether a technology could produce or magnify unequal outcomes, and if so how to mitigate this.

Key themes: combating algorithmic bias, equitable treatment, consistency, shared benefits, shared prosperity, fair decision outcomes

- › Diversity, non-discrimination, and fairness: Avoidance and reduction of bias, ensuring fairness and avoidance of discrimination, and inclusive stakeholder engagement
- › Because bias can be found at all levels of the AI and data analytics systems (datasets, algorithms, or users' interpretation), it is vital that this is identified and removed. Systems should be deployed and used with an inclusionary, fair, and non-discriminatory agenda. Requiring the developers to include people from diverse backgrounds (e.g., different ethnicities, genders, disabilities, ideologies, and belief systems), stakeholder engagement, and diversity analysis reports and product testing, are ways to include diverse views in these systems.

Areas of focus:

- › **Avoidance of unfair bias:** Take care to ensure that data sets used by AI systems do not suffer from the inclusion of inadvertent historic bias, incompleteness and bad governance models.

- › **Accessibility and universal design:** systems should be user-centric and designed in a way that allows all people to use solutions and services, regardless of their age, gender, abilities or characteristics.
- › **Society and democracy:** the impact of the system on institutions, democracy and society at large should be considered
- › **Auditability:** the enablement of the assessment of algorithms, data and design processes.
- › **Minimisation and reporting of negative impacts:** measures should be taken to identify, assess, document, minimise and respond to potential negative impacts of AI systems
- › **Trade-offs:** when trade-offs between requirements are necessary, a process should be put in place to explicitly acknowledge the trade-off, and evaluate it transparently
- › **Redress:** mechanism should be in place to respond when things go wrong.

Justice scrutiny questions

How can we ensure DDaT can help promote equitable outcomes?

- › **Remember:** emerging approaches need to ensure that they mitigate the risk of producing or magnifying disparate impacts, resulting in inequitable outcomes that make some people and communities better off and others worse off.
- › **This means:** the ethical risk that solutions might induce disparate impacts needs to be actively considered; these impacts should be anticipated, actively monitored for, carefully examined and mitigated, in order to enable acceptable ethical and equitable outcomes

How do we factor in the “bigger picture” to ensure just outcomes?

- › **Remember:** always keep in mind the wider context in which emerging approaches exist and the purpose intended, as well as considering the direction in which the technology we introduce today may head in the future.
- › **This means:** our operational and design considerations should never be isolated from the “bigger picture” of social and technological ecosystems that encompass factors and risks that are not always under our control.

Attribute 5: explicability = operate transparently

Be ready to explain a system’s working as well as its outputs. Make all stages of the implementation process open to public and community scrutiny.

Key themes: intelligibility, transparency, trustworthiness and accountability

- › **Transparency:** including traceability, explainability and communication – as Smart Information Systems can be involved in high-stakes decision-making, it is important to understand how the system achieves its decisions. Transparency, and concepts such as explainability, explicability, and traceability relate to the importance of having (or being able to gain) information about a system (transparency), and being able to understand or explain a system and why it behaves as it does (explainability).
- › **Accountability:** auditability, minimisation and reporting of negative impact, internal and external governance frameworks, redress, and human oversight. Given that Smart Information Systems act like agents in the world, it is important that someone is accountable for the systems’ actions. Furthermore, an individual must be able to receive adequate compensation in the case of harm from a system (redress). We must be able to evaluate the

system, especially in the situation of a bad outcome (audibility). There must also be processes in place for minimisation and reporting of negative impacts, with internal and external governance frameworks (e.g., whistleblowing), and human oversight.

Areas of focus:

- › Traceability: the data sets and the processes that yield the AI system's decision should be documented
- › Explainability: the ability to explain both the technical processes of an AI system and the related human decisions
- › Interpretability: helping to give users confidence in AI systems, safeguarding against bias, meeting regulatory standards or policy requirements and overall improving system design
- › System accountability: Any system, and those who design it, should be accountable for the design and impact of the system. As a minimum this should include that you can:
 - Ensure that systems with significant impact are designed to be auditable;
 - Ensure that negative impacts are minimised and reported;
 - Ensure internal and external governance frameworks;
 - Ensure redress in cases where the system has significant impact on stakeholders;
 - Ensure human oversight when there is a substantial risk of harm to human values.

Explicability scrutiny questions

How do we establish clear governance around ethical responsibility and lines of accountability?

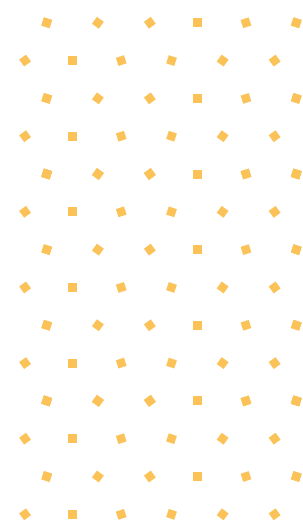
- › **Remember:** effective ethical responsibility for DDaT approaches and solutions can only be sustained across an organisation through establishing clear

chains of responsibility and accountability to assure that everyone involved focuses on their ethical duty of care towards both outcomes and the wider community.

- › **This means:** clearly define what leaders, policy makers and practitioners need to do to deliver ethical accountability in practice and who is responsible for each aspect of ethical risk management and prevention of harm in each of the relevant areas of risk laden activity (data collection, use, security, analysis, disclosure, etc.).

How do we uphold the values of Transparency and Trustworthiness?

- › **Remember:** central to maintaining public confidence in emerging approaches and solutions is respect for transparency, autonomy and trustworthiness.
- › **This means:** ensure that as far as practically possible the design, development, deployment and delivery processes of DDaT outcomes are open to public and community scrutiny, offer clear routes of redress and are mindful of people's autonomous right to choose.



Conclusion

As we have seen, the unprecedented rise of smart information systems increasingly impacts on ethics and human rights issues. AI and robotics are already widely adopted, but largely go unnoticed. Coupled with the introduction of these technologies is the ability to leverage better insights from the burgeoning amounts of data that they generate. Our responsibility in the public sector should be to prepare for the changes that result from the use and application of these technologies, using data in a more meaningful way and ensuring that we are embracing these capabilities to make improvements in public service delivery and outcomes.

As a consequence, the need to understand and promote the ethical use of emerging technologies and the data they generate and store has never been more important. We need to embed digitally ethical resilience at the very heart of our response to these unprecedented digitally-driven opportunities and challenges.

We need to improve our digital capabilities internally and grow our workforce competencies. We also need to prepare and redesign our services, so that the whole of society can benefit with better outcomes from these revised practices and ways of working. Above all, we need to design, develop, deploy and deliver place-based solutions and services that are built around core values of doing good not harm, human rights, justice, fairness, trust and transparency.

In support of these responses, Socitm is developing a digital ethics practice collection on its [resource hub](#). The collection will be built around the five-core attributes of ethical practice – Beneficence, Non-Maleficence, Autonomy, Justice and Explicability – to help leaders, policy-makers and practitioners bridge the divide between theory and practice and embed digital ethics at operational and place-based levels as follows:

- Increasing awareness amongst political and executive leadership in organisations to encourage them to dedicate more resources to identifying and addressing relevant digital ethics concerns across not just their organisations, but across the wider local place-based communities they serve.
- Educating policymakers and legal/policy teams on how to build governance structures that incentivise organisations to address digital ethics issues and proactively address the needs of local people and communities in inclusive and just ways.
- Educating technology practitioners, data analysts, and developers on techniques to address, detect, and mitigate ethical concerns with their AI and Data Analytics project implementations and solution deployments, both organisationally and at a local community level.

References

- ¹ UK guidance landscape: bit.ly/3mmc4Lv
- ² Ethics of technology, Wikipedia: bit.ly/3c1mqw
- ³ Digital ethics guide for professionals: bit.ly/2FyKolW
- ⁴ OECD AI Policy Observatory: oecd.ai
- ⁵ Ethics guidelines for trustworthy AI, European Commission: bit.ly/3msgsbW
- ⁶ Shaping the ethical dimensions of smart information systems, SHERPA project: project-sherpa.eu
- ⁷ Declaration on ethics and data protection in artificial intelligence, ICDPPC: bit.ly/3hBuBjh
- ⁸ Planting the flag, Socitm: bit.ly/planting-the-flag
- ⁹ Introducing the Amsterdam City Doughnut, Kate Raworth: bit.ly/3hEHRUa
- ¹⁰ See from what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices: bit.ly/3kkzajG

Appendix 1: emerging ethical standards

From a standards perspective the [Institute of Electrical and Electronics Engineers](https://www.ieee.org) (IEEE) is facilitating the global consultation on the development of a comprehensive set of standards for 'Ethically Aligned Design' smart information systems/Autonomous and Intelligent Systems.

Positioning 'human well-being' as a central precept, the IEEE initiative explicitly seeks to reposition robotics and AI as technologies for improving the human condition rather than simply vehicles for economic growth.

The IEEE work is focussed around the following five areas for investigation (see below) which aim to educate, train and empower AI/robot stakeholders to 'prioritise ethical considerations so that these technologies are advanced for the benefit of humanity'.

- › **Human rights:** how can we ensure that A/IS do not infringe upon human rights?
- › **Prioritizing wellbeing:** traditional metrics of prosperity do not take into account the full effect of A/IS technologies on human well-being.
- › **Accountability:** how can we assure that designers, manufacturers, owners, and operators of A/IS are responsible and accountable?
- › **Transparency:** how can we ensure that A/IS are transparent?
- › **A/IS technology misuse and awareness of it:** how can we extend the benefits and minimize the risks of A/IS technology being misused?

There are currently 14 IEEE standards working groups working on drafting so-called 'human' standards that have implications for artificial intelligence (table on following page).

Standard	Aims/objectives
<p>IEEE P7000™ - Model Process for Addressing Ethical Concerns During System Design</p>	<p>To establish a process for ethical design of autonomous and intelligent systems.</p>
<p>IEEE P7001™ - Transparency of Autonomous Systems</p>	<p>To ensure the transparency of autonomous systems to a range of stakeholders. It specifically will address:</p> <ul style="list-style-type: none"> • Users: ensuring users understand what the system does and why, with the intention of building trust; • Validation and certification: ensuring the system is subject to scrutiny; • Accidents: enabling accident investigators to undertake investigation; • Lawyers and expert witnesses: ensuring that, following an accident, these groups are able to give evidence; • Disruptive technology (e.g. driverless cars): enabling the public to assess technology (and, if appropriate, build confidence).
<p>IEEE P7002™ - Data Privacy Process</p>	<p>To establish standards for the ethical use of personal data in software engineering processes. It will develop and describe privacy impact assessments (PIA) that can be used to identify the need for, and effectiveness of, privacy control measures. It will also provide checklists for those developing software that uses personal information.</p>
<p>IEEE P7003™ - Algorithmic Bias Considerations</p>	<p>To help algorithm developers make explicit the ways in which they have sought to eliminate or minimise the risk of bias in their products. This will address the use of overly subjective information and help developers ensure they are compliant with legislation regarding protected characteristics (e.g. race, gender).</p> <p>It is likely to include:</p> <ul style="list-style-type: none"> • Benchmarking processes for the selection of data sets; • Guidelines on communicating the boundaries for which the algorithm has been designed and validated (guarding against unintended consequences of unexpected uses); • Strategies to avoid incorrect interpretation of system outputs by users.
<p>IEEE P7004™ - Standard on Child and Student Data Governance</p>	<p>Specifically aimed at educational institutions, this will provide guidance on accessing, collecting, storing, using, sharing and destroying child/student data.</p>
<p>IEEE P7005™ - Standard for Transparent Employer Data Governance</p>	<p>Similar to P7004, but aimed at employers.</p>

Standard	Aims/objectives
IEEE P7006™ - Standard for Personal Data Artificial Intelligence (AI) Agent	Describes the technical elements required to create and grant access to personalised AI. It will enable individuals to safely organise and share their personal information at a machine-readable level, and enable personalised AI to act as a proxy for machine-to-machine decisions.
IEEE P7007™ - Ontological Standard for Ethically Driven Robotics and Automation Systems	This standard brings together engineering and philosophy to ensure that user well-being is considered throughout the product life cycle. It intends to identify ways to maximise benefits and minimise negative impacts, and will also consider the ways in which communication can be clear between diverse communities.
IEEE P7008™ - Standard for Ethically Driven Nudging for Robotic, Intelligent, and Automation Systems	Drawing on 'nudge theory', this standard seeks to delineate current or potential nudges that robots or autonomous systems might undertake. It recognises that nudges can be used for a range of reasons, but that they seek to affect the recipient emotionally, change behaviours and can be manipulative, and seeks to elaborate methodologies for ethical design of AI using nudge.
IEEE P7009™ - Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems	To create effective methodologies for the development and implementation of robust, transparent and accountable fail-safe mechanisms. It will address methods for measuring and testing a system's ability to fail safely.
IEEE P7010™ - Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems	To establish a baseline for metrics used to assess well-being factors that could be affected by autonomous systems, and for how human well-being could proactively be improved.
IEEE P7011™ - Standard for the Process of Identifying and Rating the Trustworthiness of News Sources	Focusing on news information, this standard sets out to standardise the processes for assessing the factual accuracy of news stories. It will be used to produce a 'trustfulness' score. This standard seeks to address the negative effects of unchecked 'fake' news, and is designed to restore trust in news purveyors.
IEEE P7012™ - Standard for Machine Readable Personal Privacy Terms	To establish how privacy terms are presented and how they could be read and accepted by machines.
IEEE P7013™ - Inclusion and Application Standards for Automated Facial Analysis Technology	To provide guidelines on the data used in facial recognition, the requirements for diversity, and benchmarking of applications and situations in which facial recognition should not be used.
IEEE P7014™ - Standard for Ethical considerations in Emulated Empathy in Autonomous and Intelligent Systems	This standard defines a model for ethical considerations and practices in the design, creation and use of empathic technology, incorporating systems that have the capacity to identify, quantify, respond to, or simulate affective states, such as emotions and cognitive states. This includes coverage of 'affective computing', 'emotion Artificial Intelligence' and related fields.

Appendix 2: emerging guidance

The UK is playing a leading role in the development of digital ethics. This section collates the various existing ethical principles for data and AI, developed by government, public sector bodies and organisations.

Resource	Creator	Description	Who is it for?
Data Ethics Framework	Department for Digital, Culture, Media and Sport (DCMS) and Government Digital Service (GDS)	The Data Ethics Framework guides appropriate and responsible data use in government and the wider public sector. It helps public servants understand ethical considerations, address these within their projects, and encourages responsible innovation.	For anyone working directly or indirectly with data in the public sector, including data practitioners (statisticians, analysts and data scientists), policymakers, operational staff and those helping produce data-informed insight, to ensure the highest ethical standard of their projects.
National Data Strategy	Department for Digital, Culture, Media and Sport (DCMS)	The National Data Strategy (NDS) is an ambitious, pro-growth strategy that aims to drive the UK in building a world-leading data economy while ensuring public trust in data use.	For anyone working directly or indirectly with data in the public sector.
A guide to using AI in the public sector	Office for AI, Government Digital Service and the Alan Turing Institute	The guide to using AI in the public sector is a guide on how to choose, build and use AI in the public sector to ensure that government is maximising the benefits of such technology.	For public servants deciding whether AI is the right solution in their project and looking for information on how to use AI safely and ethically.
AI procurement guidelines	Office for AI and World Economic Forum	The AI procurement guidelines are a guide that addresses the ethical uncertainty around AI procurement so that officials feel equipped to use innovative technology whilst being able to mitigate risks.	For public servants and commercial specialists procuring AI. The guidelines set a baseline for the harmonisation of AI procurement standards.

Resource	Creator	Description	Who is it for?
UKSA Self-assessment	UK Statistics Authority	The UK Statistical Authority's self-assessment aims to offer researchers an easy-to-use framework to review the ethics of their projects throughout the research cycle. The self-assessment provides a timely means to identify ethical issues and shape future discussions. The process aims to support an accurate and consistent estimation of the ethical risk of research proposals.	For public servants creating statistical research projects.
Code of conduct for data-driven health and care technology	Department of Health and Social Care (DHSC)	The code of conduct is a guide that tackles a number of emerging ethical challenges associated with the use of data-driven technologies in the NHS and the wider health and care system. It applies data ethics to the specific sensitivities and challenges of the health care system. The code of conduct aims to ensure that the NHS capitalises on the possible technical benefits for users and staff.	For public servants involved in health-care specific projects who are considering the use of data-driven technologies in their work.
NHSX: A Buyer's Checklist for AI in Health and Care	NHSX	The Buyer's Checklist for AI in Health and Social Care is a short reference checklist to assist the decision-making of those responsible for procuring AI solutions in their organisations.	For procurement officials working in health and social care; for chief finance officers, service transformation and commissioning leads, in-house data scientists and analysts and other chief officers considering procurement of AI solutions.
Intelligent security tools	National Cyber Security Centre (NCSC)	The NCSC's intelligent security tools provide guidance on using AI tools in the security sector. Consideration of intelligent security tools is divided into four sections, each with a series of questions designed to aid users.	For public servants looking to use an off-the-shelf security tool that employs AI as a core component. It may also be of use to those developing in-house AI security tools or when considering AI for some non-security business function.

Resource	Creator	Description	Who is it for?
<u>The Dstl Biscuit Book</u>	Defence, Science and Technology Laboratory (DSTL)	The Biscuit Book explains some of the most common terminology in artificial intelligence and data science in an informative and accessible way. The guide is arranged as a series of easily digestible chunks that each cover a topic in a manner that provides the essential information without being too technical.	For public servants looking for a non-technical, easily-digestible guide to AI, data science and machine learning.
<u>The Magenta Book</u>	Her Majesty's Treasury (HMT)	The Magenta Book provides guidance on evaluation in government. It focuses on the scoping, design, conduct, use and dissemination of evaluations as well as the capabilities required of government evaluators. It provides guidance on how to incorporate evaluation through the design, implementation, delivery and review stages of policy making. It also explains how results can be interpreted and presented, and what should be considered in this process.	For the policy, delivery and analytical professions, all of which are responsible for securing and using good evidence.
<u>Analytical Quality Assurance Book (AQuA)</u>	Her Majesty's Treasury (HMT)	The Aqua Book is a good practice guide on analysis and analytical models. It was created by a cross-departmental working group on analytical quality assurance.	For those working with analysis and analytical models.
<u>Service Standard - Service Manual</u>	Government Digital Service (GDS)	The Service Standard is a 14-point manual to help teams to create and run great public services. The Service Manual focuses on user needs, as well as the team involved in the project and how to create the service.	For public servants creating public services using data and AI.

Resource	Creator	Description	Who is it for?
Digital Ethics Charter	Midlands Accord and East Accord	The Digital Ethics Charter is a set of common principles that digital professionals and those working with “technology for public use” can adhere to.	Professionals working in health and social care
Data Ethics Canvas	Open Data Institute	The Data Ethics Canvas helps identify and manage ethical issues – at the start of a project that uses data, and throughout. It encourages you to ask important questions about projects that use data, and reflect on the responses.	For anyone working directly or indirectly with data in the public sector, including data practitioners (statisticians, analysts and data scientists), policymakers, operational staff and those helping produce data-informed insight, to ensure the highest ethical standard of their projects.
Understanding artificial intelligence ethics and safety	Alan Turing Institute	A guide for the responsible design and implementation of AI systems in the public sector	For everyone involved in the design, production, and deployment of a public sector AI project: from data scientists and data engineers to domain experts, delivery managers and departmental leads.
Digital Ethics Review	techUK	Policy paper which aims to help ensure that digital ethics is regarded as relevant and beneficial to the real lives people lead.	Leaders, policy makers and practitioners across the public and private sector
Decision-making in the Age of the Algorithm	NESTA	A guide for public sector organisations on how to introduce Artificial Intelligence tools so that they are embraced and used wisely by practitioners.	For those working with analysis and analytical models.
Applied AI - ethics typology	Digital Catapult	Supporting the application of AI ethics through a typology of principles, tools and resources and services.	To help AI developers build value-aligned products

Resource	Creator	Description	Who is it for?
Guidance on AI and Data-Protection	Information Commissioner's Office	Guidance to help organisations mitigate the risks specifically arising from a data protection perspective, explaining how data protection principles apply to AI projects without losing sight of the benefits such projects can deliver.	For everyone involved in the design, production, and deployment of a public sector AI and Big Data projects: from data scientists and data engineers to domain experts, delivery managers and departmental leads.
Artificial Intelligence and Public Standards	Committee on Standards in Public Life	CSPL report and recommendations to government to ensure that high standards of conduct are upheld as technologically assisted decision making is adopted more widely across the public sector.	Leaders, policy makers and practitioners across central government and the wider public and private sectors.
Guidebook of data literacy tools	Data Justice Lab	The guide provides an overview of different types of tools that aim at educating citizens about datafication and its social consequences.	For anyone working directly or indirectly with data in the public sector, including data practitioners (statisticians, analysts and data scientists), policymakers, operational staff and those helping produce data-informed insight, to ensure the highest ethical standard of their projects.
Consequence Scanning Manual	Doteveryone.org.uk	It introduces Consequence Scanning – an iterative development tool to help organisations think about the potential impact of their solutions or services on people and society.	For anyone directly or indirectly involved with the design of public sector digital and data solutions or services.
A Guide for Ethical Data Science	Royal Statistical Society (RSS) and Institute and Faculty of Actuaries (IFoA)	It is intended to complement existing ethical and professional guidance and is aimed at addressing the ethical and professional challenges of working in a data science setting.	For professionals and officials working in the area of data science.

Resource	Creator	Description	Who is it for?
Creating City Portraits	Doughnut Economics Action Lab	The guide sets out the DEAL City Portrait methodology to all who are interested in downscaling the Doughnut to their city or place	For anyone directly or indirectly involved with the design of civic sector digital and data solutions or services.
Building Digitally Inclusive Communities	International City/County Management Association and partners	Model framework for people, businesses, and institutions will have access to digital content and technologies that enable them to create and support healthy, prosperous, and cohesive 21st-century communities.	Leaders, policy makers and practitioners across central government and the wider public and private sectors.
ICO Accountability Framework	The Information Commissioner's Office (ICO)	The framework is divided into 10 categories and contains expectations and examples of how your organisation can demonstrate your accountability.	Leaders, policy makers and practitioners – accountability is one of the key principles in data protection law – it makes you responsible for complying with the legislation and says that you must be able to demonstrate your compliance.

About this briefing

Author

William Barker - Associate director and practice lead on digital ethics

Editor

Martin Ferguson - Director of policy and research

Designers

Magdalena Werner - Senior creative designer

Benjamin Hughes - Graphic designer

Have your say

We always welcome feedback and discussion on the contents of our publications.

Martin Ferguson

Director of policy and research
martin.ferguson@socitm.net

Nadira Hussain

Director of leadership development and research
nadira.hussain@socitm.net

Get in touch

Website: www.socitm.net

Email: inform@socitm.net

Tel: 01604 709456



Join the conversation...



@Socitm



Socitm