Delivering Artificial Intelligence

Striking the right balance

Artificial Intelligence (AI) for good

Maslow is widely known for his hierarchy of needs theory. Humans need to satisfy their physiological needs in the first instance, food, water, shelter, clothing, sleep, etc. before they can get to an ultimate point of self-actualisation, creativity, experience purpose, meaning and inner purpose.

Whilst not all will necessarily get to this point the other levels such as safety and security, love and belonging and self-esteem (not necessarily in this order) all are levels we as humans aspire to.

AI is widely expected to help humankind by improving productivity, improving customer experiences through automation and intelligent processing of increasingly complex tasks. The potential for AI to detect patterns and impacts of events (that are too complex and expensive to manually detect) opens the doors to unlimited potential to solve complex problems for the betterment of humans and society. Traditionally AI responds to inputs and creates an output based on certain rules (algorithms) but next-generation AI, such as generative, takes this to the next level by creating new data that is not pre-defined.

This can be complex text, images, video and audio, for every positive use case a negative connotation can be imagined.

"If the only tool you have is a hammer, it is tempting to treat every problem as if it were a nail" – another of Maslow's observations. AI turns this on its head, with the advent of mass consumption of generative AI through apps and foundation models such as ChatGPT, AI is the potential answer to all problems… or is it? Of course not. AI is already part of our everyday lives. It is no more remarkable than a toaster was for generation X. Innovation throughout time has always been about something new that adds value. AI is a powerful tool that can form part of a solution, it has always been constrained by the available compute processing power. With the advent of more powerful computing, using AI in a responsible, ethical and transparent manner is the challenge.

It is against this backdrop that CGI and the GSA held a roundtable discussion to explore the readiness of organisations to use AI ethically, responsibly and with the accountability of service providers and buyers using the correct frameworks for AI-enabled programmes. The diverse group of attendees included CEOs, industry experts, academics, analysts and thought leaders. The following report summarises the themes explored on the day.

Executive Summary

In partnership with the GSA, CGI is embarking on research to better understand how responsible AI can be delivered with the correct controls in place for suppliers and buyers alike.

Through open dialogue and consultation with policymakers, industry experts and academia we hope this first paper will increase awareness and stimulate discussion and debate. Our intention is to build on the themes identified and provide guidance to the sourcing market regarding buyers' concerns and challenges in procuring AI solutions and the risk suppliers face in protecting their Intellectual Property.

The discussion centred around the following questions:

1. How ready are organisations versus public sector departments to use AI responsibly, ethically and sustainably?

2. How fit for purpose are the current AI frameworks for accountability?

3. How do we ensure accountability and responsibility between buyer and service provider in the delivery of AI-enabled programs?

The inaugural round table took place on 26 March 2024 under Chatham House rules. The combined wisdom in the room identified several themes that are prevalent and should be front of mind when it comes to future sourcing of AI.

The following themes surfaced during the conversation:

Clarity
Being clear on the intended use, the data needed for training and data provided and generated through AI outputs, risks this might pose, benefits and current state

Use
Don't jump to edge use cases to create an environment that becomes too risk-averse or oversimplify introducing unnecessary risk

Balance
Managing the solutioning, procurement, contracting and readiness of the business in an ethical way

Business change and pace
Managing the desire to do something with the time needed to do it effectively and efficiently

Metrics
Tangible ways of being able to measure success with an agreed baseline

The next section summarises the conversation that took place exploring the themes described above. Finally, a summary of the roundtable will conclude this paper and explore possible next steps.

Something old or something new?

Whilst tempting to think of the release of ChatGPT on 30 November 2022 as a pivotal moment in the advancement of AI, we must go back to the 18th Century to understand where the thinking started. Nicolas de Condorcet, a mathematician, formulated theories in his publication Sketch in 1794 that envisioned future society. His central tenet being that advancement in the sciences would lead to individual freedom, moral compassion and material affluence. Clearly, the computer had not been invented but the argument that "new instruments, machines… can improve quality…accuracy of production and can diminish the time and labour expended on them" was made.

Whilst many organisations are wrangling with the theoretical implications of pursuing a particular AI strategy, one of the attendees made the crucial point that the question to ask when starting an AI journey is: what is the intended use? This observation helped to debunk the fears that can manifest themselves when using edge use cases as examples. Whilst the fear of technical singularity is in the minds of those who worry we will end up in a human zoo, the reality is we are currently living in a world where narrow AI exists and is used by all of us in our everyday lives. As AI evolves it will seamlessly integrate as our consumption of social media changes and work productivity through co-pilot tooling and whisperer advances (speech to text).

What is the problem you are trying to solve – avoiding AI is the solution but what is the question?

Thinking back to Maslow, be clear what it is you are trying to achieve. Don't just jump to the edge cases that present risk and give a reason not to do something. The regulatory frameworks and guard rails are not there to stifle innovation but to promote a 'yes if' culture rather than a 'no because' mindset. Use cases for generative AI should be proportionate. If you are developing an assistant to help patients find the right car park at a hospital that's minimal risk, but if you are trying to develop something with a medical impact like intubating a patient then that is high-risk, not so much a human in the loop as human only. Starting with a known and clear need and scaling the solution to new potential and functionality is a best practice approach to both technology use, change management and greater adoption.
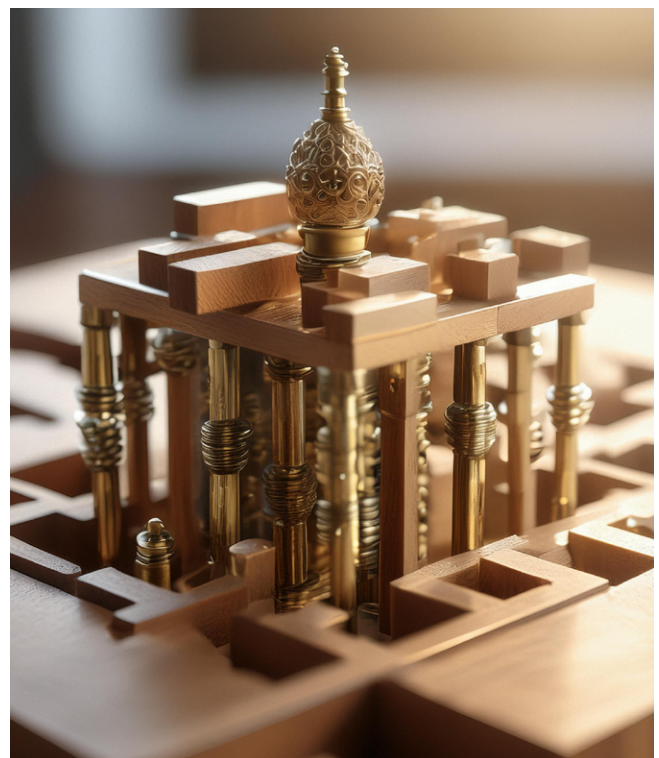
What is the problem you are trying to avoid?

Stay off the front page and out of the courts. Whether the aim is to increase productivity in a particular area or improve content creation and report writing there is a balance between inward focussed initiatives and customer-centric external interactions. Whilst it is tempting to be enthused by the claims of removing the equivalent of seven hundred agents by a financial services organisation, we need to balance this with the viral coverage a well-known courier received when their bot authored a poem about how bad they were. If you start with a clear question or problem statement the data and model will become clearer and will be fit to the purpose intended.

Some are more ready than others

The consensus was that some parts of organisations are ready but without a consistent definition of governance and appropriate security and privacy concerns boxed off, there is considerable variability in readiness and understanding. The implications of how to adopt AI effectively moving forward as part of everyday business DNA is not necessarily understood - as highlighted by the number of pilots underway versus those in production. Recent analysis identified several pain points but fundamentally incremental, tactical non-integrated approaches that do not release the full value, but then create processing challenges further down the line. One of the participants highlighted further research that showed 73% of CIOs said the lack of success metrics was their biggest concern.

The implications of how to adopt AI effectively moving forward as part of everyday business DNA is not necessarily understood.

Beware...edge use cases...

During the discussion, it was clear there was a tendency to look at the edge use cases as the ones to address regarding putting them into risk categories or regulatory principles. Working through a predictive policing hotspot algorithm is different from changing your address details with the bank. The fundamental questions when solutioning or sourcing AI is no different to procuring Java development.

How ready are organisations? Understanding the balance between what is meant by the broad definition of AI and what the implications are for data security and privacy, business change and the need to address data and technical issues means there are numerous plates spinning. The fact the definition is so broad means that we should take care not to talk ourselves out of potential applications.

This also goes to the heart of the contracting challenges. If a CIO or procurement team need to balance the risk of paying compensation or losing customers against a more traditional cost-saving productivity improvement, such as an outsourcing initiative, they may stick with what they know. The AI debate is like those had 25 years ago about outsourcing non-core business for cost and quality efficiencies, indeed the logic – cost reduction, improved quality and freeing up more time for innovation still holds true.

Don't let perfect be the enemy of good – the balance beam example:

So how to traverse the balance beam, internal versus external? Of the numerous pilots underway the collective experience in the room was there is more of a focus on improving agent productivity in the Customer Experience (CX) space over customer interaction. The sentiment is one of caution and lockdown of data, so any generative training models focus internally in a secure ringfenced way. So do organisations remove, replace or augment agent/customer assist in this context?

An observation shared was the tactical non-integrated approach some companies are taking to the development of AI when the advantage lies in a more holistic integrated approach. An example discussed is in healthcare - genAI offering the option to look at non-surgical solutions to a particular problem when interrogating a certain dataset. The potential is there, but people are apprehensive, and it is a question of finding the right balance in AI for good. And how does this happen in a careful thoughtful way?
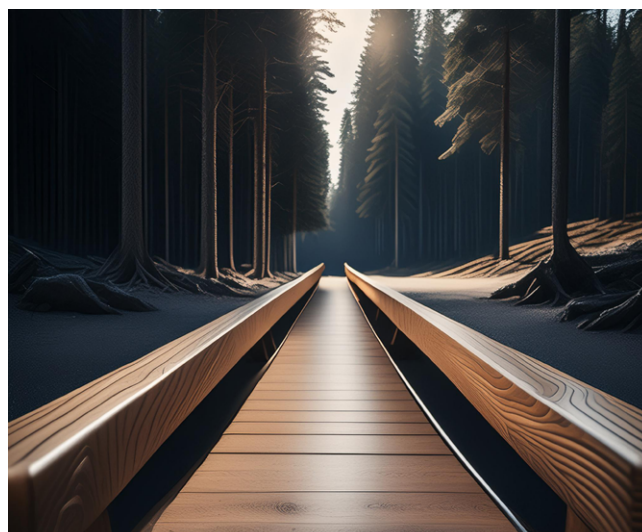
Framework flooding

Legislative

The EU Act, which like GDPR may become a de-facto standard is the first in the world to be ratified.

It has four risk categories for AI solutions: unacceptable, high, limited and minimal. The transparency requirements for generative AI will also be assessed throughout the lifecycle of a solution. The sourcing industry needs to navigate these, and the inter-country direction taken which could be different in each offshore location. The approach in the UK outlined in the Bletchley Declaration in October 2023, the regulatory principles contained in the current documentation include safety, security and robustness, appropriate transparency and explainability, fairness, accountability and governance, contestability and redress. Further discussions are due to take place in 2024.

UNESCO (United Nations Educational, Scientific and Cultural Organisation) have developed two framework methodologies that support countries in implementing their recommendations on AI ethics. The Readiness Assessment Methodology (RAM) and Ethical Impact Assessment are ones to watch, with a cohort of 50 countries using these to create an evidence base. Some countries involved have not published an AI strategy so are bucking the trend and starting with the ethical guidelines first, then the AI strategy. One participant explained how their corporate approach was to tailor the AI implementation lifecycle to the Learning Health model which uses an evidence-based statistical approach to proving certain problem hypotheses by engaging stakeholders to help drive change to transform.

Procurement

Crown Commercial Services have taken a collaborative approach to developing the next iteration of the RM6200 framework, which ends in November 2024. This was welcomed and seen as best practice, with classic market sounding. The fundamental requirements cannot simply be imposed because of the nuances in the way AI models work and the speed at which they are being developed and deployed.

For the public sector in the UK, the additions to the Data Protection Act around requirements for identity protection for citizens will no doubt filter through to the commercial sector. The challenge across all sectors is akin to the early days of sourcing, with the analogy of a balance beam being appropriate. If a contract is too narrow it stifles innovation - too wide, and there is too much room for interpretation. Given the complexities of modern procurement and the multitude of stakeholders involved, there is a real danger that the risks will appear insurmountable.

The ultimate responsibility is one area that is proving challenging when advising buyers and suppliers. If the training data is the clients, the AI stack coming from a vendor with a service wrap from a systems integrator; who owns the solution and therefore the risk? How is accuracy defined? A vendor when selling an AI solution must ask what am I committing to? A buyer should be clear on what am I actually buying? Is it clear to all parties, with a collective understanding?

Educating all stakeholders on the responsible use of AI is fundamental to effective delivery

Implementation and delivery

Using an evidence-based approach to developing AI solutions is something that Covid-19 pandemic highlighted globally and in a very public way. This was the first time the public had been provided the data and information on a rapidly changing crisis event.

Whilst most citizens are familiar with the rapid development of the various vaccines, many fewer fully appreciated that the rapid development relied on scientific randomised control trials (RCTs) using statistical models to evidence the efficacy of vaccines. This approach was used when developing virtual assistants that would be used in some countries to help citizens triage their symptoms before requesting a test. For these to become operational, elevated levels of reliability needed to be demonstrated and underpinned with statistical validation.

Educating all stakeholders on the responsible use of AI is fundamental to effective delivery. All participants agreed transparency and communication were key. There is a suite of effective toolsets available to support the explainability challenges the more complex solutions adopt. Ensuring these are built into the overall solution design and business change implementation are seen as positive enablers.

If the data is not labelled correctly and the use of Generative AI is just internal there is a significant risk that exposure to confidential material can occur. Citing sources is a way of mitigating this when creating content. The challenge of course with generative AI is the complex nature of neural networks where developers cannot emphatically say how an output has been generated.

What's the baseline and how is success measured?

Is there a baseline in place that success criteria can be measured against? The term pilot purgatory is growing across the industry which implies that there is confusion and a sense of lack of direction.

A point well made during the discussion was whether staff are more concerned with answering a Board's question "what are we doing with AI" as opposed to developing a use case that can then have tangible benefits measurement through clearly articulated metrics. The metrics should be applied in readiness to terms of data and technology.

Metrics is a concern named by a significant number of CIOs linked to the points above. There is a developing weariness within some areas at the mention of AI because of the opaque measurement of benefits. If a chatbot can save x human hours but in fact causes an increase in complaints and potential compensation and lost business. The challenge being how to measure cause and effect. More importantly, when preparing a chairman's statement or Board presentation it's nice to know the proof of concept or pilot went well but when will the results come? Why not use a more traditional route and take a 30% reduction using offshore sourcing?

Return on Investment

Investment is different this time. The amount that needs to be invested, even when using opensource, is significant. So, before any commitment is to be made, the consensus in the room was that the expected benefit should be realised before any commitment for the big investments happens. The pain points discussed and researched may break down into fundamental observations. It is the incremental, tactical and not integrated approaches that may create more of a headache for processing further down the line.

So, when the board has the following conundrum – getting 30% cost savings by going off shore or a potential 60% using an untested recent technology that relies on data being in a good place – will the board have sufficient confidence to commit to a new technology?

It's the data - foundation and fundamentals of data maturity

AI only works with the right data foundation. Most organisations have plenty of data but lack the maturity to move this to usable information to improve knowledge and become a wiser organisation. The group discussed data journeys looking at breaking data fundamentals into the component parts, including the use of information flows and models, data ownership including best practices and business aims, and data governance. Simply, review, retain or delete with clear rules and share appropriately. With appropriate data standards agreed between suppliers and buyers and appropriate data security, including role-based access controls and secure closed applications on cloud environments, with information archives enabled to reduce risk, organisations can move to reliable data-driven decision making. This alongside regular data maturity assessments to make sure they have the right data foundations in place for the use of AI.

Of the four elements of AI (data, algorithms, compute and human) data is the foundation. It is where the risk lies, particularly accessibility. Organisations such as DAMA (Data Management Association) has a solid foundation that can be extended to AI uses of data. There is no defined step process for ensuring data adequacy, but at all points, decision-makers need to be clear on the risk they are accepting. For example, there are students now using training models and public data that are potentially infringing copyright and intellectual property without awareness. Encouragingly there is more discipline within the developer and technology community based on the quality of good governance, regulation and user understanding. Albeit in a recent survey of 200 enterprises one of the attendees highlighted that 63% felt they were ready from a data perspective and their concerns were more about technical readiness, with people readiness being lower than 50%.

Responsible use of AI (RAI)

How RAI develops is core to the future of AI. As with the iPhone, there will be a period of 'How does one fit AI into the way we work now' as opposed to 'How will AI change the way we work in the future?' Within the sourcing community on both sides of the fence, it was felt that open dialogue was the most productive way forward. Getting the strategy and governance aligned along straightforward principles, addressing the ethics, trustworthiness and explainability of new emerging technology is paramount. During the debate, these three principals were broken down and discussed further.

Firstly, ethical considerations, fairness and inclusivity throughout the AI lifecycle. How do stakeholders ensure that discrimination, bias, disinformation, and plagiarism do not manifest? How can AI be used to positively promote equality? How can this and other data be tracked to show a positive impact on an organisation's environmental, social and governance (ESG) and sustainability position? The initial thoughts include aligning to human values including safe value-based human-in/off-the-loop design, and continuous supervision of AI alignment to human objectives, goals and values.

Secondly, trustworthiness is synonymous with transparency, accountability and reliability, how an organisation empowers and oversees the creation of AI solutions. In particular, accountability at both an executive and operational level regardless of their position in the hierarchy means accepting responsibility for designing, developing, deployment and final operation of AI systems and components. Being consistently transparent, clear documentation on the provenance of data and the associated models being deployed, openly communicating compliance with the associated AI regulations by looking to exceed where possible and continuously enhancing the existing delivery frameworks.

Demonstrating the ability to reproduce results in a controlled environment and manner.

Finally, explainability meaning for the data scientists, business analysts and developers involved in the creation of AI systems make interpretability and explainability part of the core design. Embrace the toolkits, frameworks and techniques to prove statistical significance where possible. An example was used in a generative Freedom of Information application that cites all sources when producing the text.

The participants also debated the position that the organisations as a whole need to be trained in the principles, risks and regulations so there is a continuous culture of learning and training.

Human in the loop (HITL, HOTL))

Human 'in' the loop involves AI providing AI driven advice to a human, Human 'on' the loop involves ongoing monitoring and auditing of AI driven outputs and advice. The recent high-profile issues, such as the well-known courier whose bot insulted a customer to the airline that fought an expensive legal battle over a comparatively minor sum by disputing two contradictory statements (selecting the one in their favour), highlights the need for caution when developing AI solutions. Having to factor in the risk of something going wrong and compensation having to be paid is important for any organisation that wants to protect its brand. This is even more important when making decisions that could fundamentally affect a citizen. A balance between AI assisted people and people assisted AI to provide the emotional reassurance humankind might need to believe that AI will not take over the world. The chief scientific officer for the National Police Chiefs Council stated that the intention is a human will always make the decision when using AI solutions in policing. However, when or indeed if Artificial Super Intelligence ever surpasses human intelligence, it is unlikely to become uncontrollable and irreversible was the collective agreement in the room.

Conclusion

Regardless of whether an
AI solution is for internal
consumption, a client or jointly
created, the collective view was
that those involved in designing
and implementing should
be trained and aware of the
limitations involved.

By taking an integrated approach including procurement considerations, protection of IP, awareness of training model boundaries, and associated reputational risks, the AI sourcing journey need not be a difficult one. More a question of informed, reflective collaborative approaches to qualify a particular strategy (or not)!
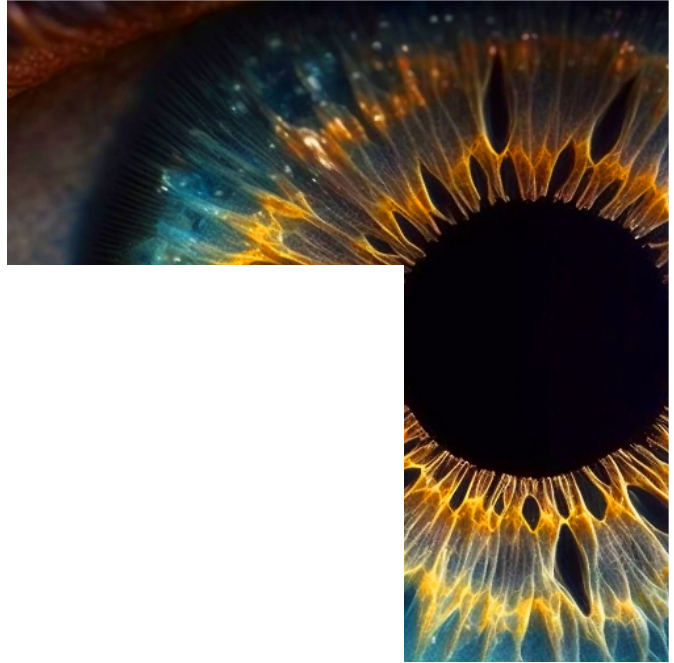
For further information please contact:

Nick Dale, Vice President, Police and Public Safety AI Lead, CGI
nick.dale@cgi.com

gsa-uk.com
admin@gsa-uk.com
LinkedIn: @GSA - Global Sourcing Association

About GSA

The Global Sourcing Association is the industry association and professional body for the global technology and business services industry. We sit at the intersection between buyers, vendors and advisors and our unique combination of expertise, insight and impartiality is why more than 10,000 members trust the GSA to represent them. Our work on best practices, standards, and professional development supports the future growth of the industry and the professionals working within it.

We are a social enterprise committed to promoting sustainable and ethical sourcing to create a positive future for our businesses and our shared planet.

About CGI

Founded in 1976, CGI is among the largest IT and business consulting services firms in the world.

We are insights-driven and outcomes-based to help accelerate returns on your investments. Across 21 industry sectors in 400 locations worldwide, our 90,000 professionals provide comprehensive, scalable and sustainable IT and business consulting services that are informed globally and delivered locally.

Our commitment: Insights you can act on.

cgi.com/uk