Reliance Cyber Artificial Intelligence Mission Statement

Artificial Intelligence is a disruptive technology that is having a large impact on the global digital landscape. At Reliance Cyber, we believe there are substantial rewards for implementing well-designed AI systems, but they need to be deployed carefully and thoroughly evaluated. The risks posed by poor use of the technology are significant, and implementation must adopt a measured approach, deploy AI services where it helps our customers and us, but we will resist use cases that are needless or introduce new risk.

Artificial Intelligence enables new capabilities that Reliance Cyber must take advantage of. AI offers opportunities for rapid response and remediation, ingestion and analysis of previously unmanageable amounts of information, and the precision to identify critical anomalies in noisy data sets. By using these capabilities, we can offer our customers stronger and faster protection than was previously possible. Over time AI will enable new functionality and approaches to cybersecurity, and we must be flexible in assessing and deploying them where appropriate and safe.

Artificial Intelligence will be used to support our staff, not replace them. We believe our customers value the relationships they have with our cyber security experts, and we have no desire to water down these relationships. Artificial Intelligence will be used to support staff, augmenting but not automating their human intuition. Our 'human-in-the-loop' designs must ensure that AI handles the scale, while our staff handle the strategy. We will use these tools to make more information available, enable quicker reactions, and allow our staff to deploy their expertise where it is needed the most. While AI may automate low-impact actions, human authorisation remains our default requirement unless the risk of misclassification is demonstrably low.

Artificial Intelligence must only be used where we observe genuine benefits. We are aware that AI automation is being needlessly built into many digital services, and we will not follow this trend. All our AI tools are deployed cautiously, noting the risks of non-determinism in security processes and utilising extensive logging to allow for their continual assessment. This allows our Research and Engineering teams to carry out quantitative and qualitative research to evaluate the success and efficacy of our implementations, ensuring our services remain beneficial to our customers. In cases where AI automation does not offer the benefits we had hoped for, we will reassess, aim to improve, or ultimately be prepared to remove any functionality that does not achieve the desired results. We will aim to make as much of this data public as possible, while respecting the security and privacy of our customers.

Our Artificial Intelligence systems must be built to be resistant to cyber-attacks. As well as new opportunities, AI enables new types of cyber-attacks that we must actively counter, including both attacks against our customers as well as attacks against our AI systems themselves. We must incorporate rapidly developing and researched defence measures to ensure our platforms resist these attacks. This must be confirmed by testing and red teaming, so that we are certain that anything we implement does not increase our customers' attack surface.

Our Artificial Intelligence services must adhere to UK data sovereignty and data protection requirements. AI introduces new risks around data sovereignty and protection, and we must prioritise these over the latest model releases. This means we will not route our customers' data to models or data centres outside of the UK. Our commitment to UK-based processing ensures that our customers' regulatory requirements are never sacrificed for the sake of rapid updates. Where functionality depends on models only available outside of UK data centres, we will not enable that functionality, as our customers' security is our primary objective.

Our customers' data will not be used to train third-party models. When using third-party models, we only use licensing that precludes the use of data for model training. If, in the future, we train or fine-tune our own internal Reliance Cyber models, we will only use data with the express contractual consent and agreement of customers who wish to take part.