AI Explainability Statement

Disclaimer

Copyright © OneAdvanced 2024.

This document contains confidential and / or proprietary information. The content must not be disclosed to third parties without the prior written approval of OneAdvanced Group Limited or one of its subsidiaries as appropriate (each referred to as "OneAdvanced"). External recipients may only use the information contained in this document for the purposes of evaluation of the information and entering into discussions with OneAdvanced and for no other purpose.

Whilst OneAdvanced endeavours to ensure that the information in this document is correct and has been prepared in good faith, the information is subject to change, and no representation or warranty is given as to the accuracy or completeness of the information. OneAdvanced does not accept any responsibility or liability for errors or omissions or any liability arising out of its use by external recipients or other third parties.

No information set out or referred to in this document shall form the basis of any contract with an external recipient. Any external recipient requiring the provision of software and/or services shall be required to enter into an agreement with OneAdvanced detailing the terms applicable to the supply of such software and/or services and acknowledging that it has not relied on or been induced to enter into such an agreement by any representation or warranty, save as expressly set out in such agreement.

The software (if any) described in this document is supplied under licence and may be used or copied only in accordance with the terms of such a licence. Issue of this document does not entitle an external recipient to access or use the software described or to be granted such a licence.

The development of OneAdvanced software is continuous and the published information may not reflect the current status. Any particular release of the software may not contain all of the facilities described in this document and / or may contain facilities not described in this document.

OneAdvanced Group Limited is a company registered in England and Wales with registration number 05965280 whose registered office is at The Mailbox Level 3, 101 Wharfside Street, Birmingham, B1 1RF.

A full list of its trading subsidiaries is available at www.oneadvanced.com/privacy-policy/

OneAdvanced is committed to transparency and accountability in our use of artificial intelligence (AI). This Explainability Statement provides clear insights into how OneAdvanced AI operates, the data it uses, and how it makes decisions. This document is designed to meet the expectations outlined in the ICO's guidance on Explaining Decisions Made with AI, in collaboration with The Alan Turing Institute and the EU AI Act.

Our AI Explainability Statement provides a transparent overview of our AI processes, aiming to foster trust through clarity and understanding. Additionally, it supports our customers in meeting their legal obligations and due diligence requirements regarding transparency in AI, throughout the procurement and deployment stages.

Commitment to Responsible AI

OneAdvanced adopted the following AI principles that best reflect our values when designing, developing, and deploying OneAdvanced AI:

Transparency and explainability - We developed OneAdvanced AI in ways that are transparent and explainable to those who use or are affected by them. We provide meaningful information, appropriate to the context, to foster understanding of OneAdvanced AI, to make people aware when they are interacting with AI and, if they are affected by an automated AI decision, to understand that decision, the logic involved and the possibility of challenging that decision to a human reviewer.

Fairness and inclusivity – When developing OneAdvanced AI, we took into consideration that people must be treated fairly, with dignity, and with respect for law and their human rights. We implemented measures, consistent with the state of the art, to respect, protect and promote diversity and inclusiveness at all stages throughout the AI life cycle, so as to prevent unfair bias or discrimination against individuals or groups on the basis of race, colour, descent, gender, age, language, religion, sexual orientation political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, disability or any other grounds.

Robustness, safety and security – We developed a robust, secure, and safe OneAdvanced AI throughout its entire lifecycle so that, in conditions of normal use, foreseeable use or misuse, or other adverse conditions, it functions appropriately and does not pose unreasonable safety risk. We enabled traceability in OneAdvanced, including in relation to datasets, processes and decisions, to enable analysis of their outcomes, appropriate to context and consistent with the state of the art. We adopted risk management measures to identify and mitigate potential risks related to our OneAdvanced AI, before and after its deployment. These may include consultation with diverse stakeholder groups and domain experts to identify potential risks, including those related to privacy, security, safety, bias, harmful content and infringement of intellectual property, and mitigations we can take.

Privacy and data protection – We respect people's rights to protection of their personal data. OneAdvanced AI processes personal data in a manner that is transparent and fair, for specified purposes, and on the basis of consent or another legitimate basis laid down by applicable data protection laws. We implemented mechanisms that enable individuals to access personal data processed by OneAdvanced, and to have inaccurate or outdated personal data corrected. Where individuals are adversely affected by an automated AI decision, they must have the ability to

contest that decision to a human reviewer who will undertake a meaningful review of it, with authority to approve or reject the decision.

Accountability and Responsibility – OneAdvanced developed the OneAdvanced AI (a developer or provider under the EU AI Act), while our customers will deploy it (deployers under the EU AI Act)

We remain accountable for the proper functioning of OneAdvanced AI and for the respect of the above principles, based on our role, the context (including applicable legal and regulatory frameworks), and consistent with the state of art. We demonstrate our accountability through our governance structures, and through our actions and decision-making processes - for example, by implementing policies and procedures to manage OneAdvanced AI throughout its lifecycle.

OneAdvanced AI is deployed in a way that can be easily switched on or off. Therefore, the customers can opt out of OneAdvanced AI at any time.

Human-centric – OneAdvanced AI is a tool for humanity, not a replacement for humans. OneAdvanced AI supports human agency, augment human capabilities, and enhance human creativity.

Social well-being – OneAdvanced AI has the potential to transform the UK society for the better through our mission to power the world of work.

Environmental sustainability - OneAdvanced AI helps to drive environmental sustainability, by utilising smart technology to optimize resource use, minimise waste, and protect natural ecosystems, while reducing the carbon footprint of our operations. See our AI ESG statement.

OneAdvanced AI purpose and role of AI

Artificial Intelligence (AI) is changing the way we work. As UK organisations embrace AI to drive productivity, they face the threat of data security risks, yet they increasingly understand the need to embrace AI to stay competitive. The challenge is that organisations are worried about putting company and personal data into public AI models and are concerned about data sovereignty, privacy and security.

OneAdvanced AI is built with this in mind.

OneAdvanced AI is a private, secure AI service with an intuitive chat interface for role, department and sector specific work in UK organisations.

OneAdvanced is committed to delivering long-term value to our customers by investing in technology, processes, and data. OneAdvanced AI is **secure, trusted and hosted in the UK** to maintain UK data sovereignty for users, empowering businesses while ensuring compliance.

Our system relies on information provided by user by way of data and prompts with an LLM and machine learning algorithms to process data and provide informed recommendations.

The AI-assisted outcome is clearly demonstrated in each response of an end user's chat interaction with OneAdvanced AI.

OneAdvanced AI design, functionality and limitations

OneAdvanced AI has been designed and built to address the needs of UK organisations to power their world of work, while maintaining privacy and UK data sovereignty.

Functionality

- **Protected data using AI -** Our UK-based private hosted AI model ensures that your data remains in the UK, under your control and in compliance with the applicable data protection legislation (e.g. UK GDPR).
- **Intuitive chat interface** – for end users to simply interact within OneAdvanced AI.
- **Work privately -** Our private spaces allow you to upload files to a secure environment, bringing the power of AI to your documents without compromising on security.
- **Enhance productivity -** Our Retrieval-Augmented Generation technology combines data sources to enhance work, making it easier to complete tasks and achieve goals.
- **Fully encrypted** – Being secure by design, the end-to-end OneAdvanced AI environment is fully encrypted.
- **Identity management controlled by the organisation** - providing security and convenience for the end user accessing OneAdvanced AI.
- **Customer Data is <u>not</u> used to train AI Models** – Aligning to trust, privacy and safety, prompts with data are not stored or used to train OneAdvanced AI models.
- **Unified user experience** - Aligned to the strict standards we set for accessibility and an optimised user experience, OneAdvanced AI is WCAG 2.2 AA compliant. OneAdvanced AI is accessed via a web browser and is browser responsive, aligning to customers' preferred way of interacting with the solution.
- **Data visualisation** – create charts and graphs to support data analysis objectives to aid informed business decisions.

Use Cases

The below are examples of how we help power the world of work for UK organisations with OneAdvanced AI.

- **Query company information** - new joiner wanting to know how much they can spend on lunch whilst at a client visit – asks the OneAdvanced AI to advise based using the company policy they have uploaded.
- **"I need help"** - A user of OneAdvanced AI needs to run an ideas workshop with customers to get their input on product development – how would you suggest I structure the session? Here is my product strategy document please use this as reference and supply your response in a matter-of-fact tone. The workshop will run for 4 hours.
- **Comparing two documents** – An organisation has a current 3 year-contract with a supplier that is coming up for renewal and the supplier has advised their T&Cs to have changed – we need to compare our original contract with the new one to see what has altered.
- **Summation** - I have just launched a new service for my customers and the latest sales numbers are in. Can you summarise this data (Excel) and provide insight along with a chart that displays the best performing sectors. Suggest areas to focus on for improvement.

When OneAdvanced AI should not be used

OneAdvanced AI should not be used for the prohibited uses outlined within the [OneAdvanced Acceptable Use Policy](#)

OneAdvanced AI High Level Architecture
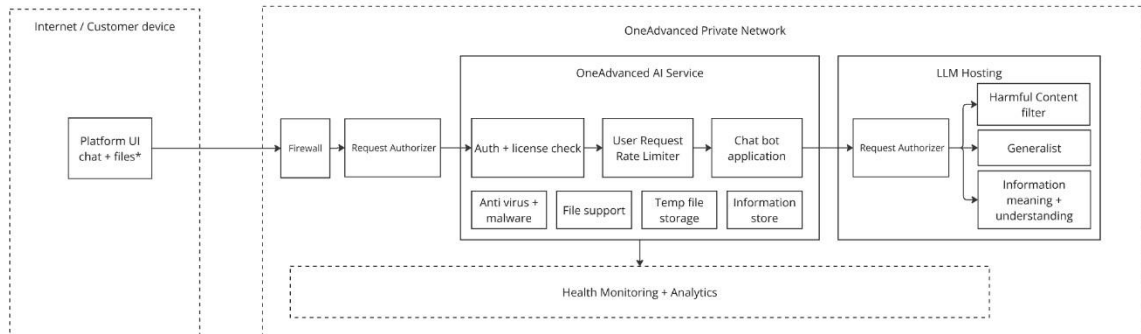
All network communications are over encrypted | TLS 1.2 or better
All data is encrypted at rest
Each users information is segmented from other users
No user prompts or contents of their data is logged
Only health data to ensure the system is operating effectively is stored
Files are all scanned for malware / viruses and deleted if they are infected
Files are only kept up to the end of the session and are removed
Zero trust approach to authorize requests at each service boundary

Data use and sources

AI Model

OneAdvanced AI leverages an open-source large language model (LLM), hosted privately in the OneAdvanced Platform in the UK.

We tested a shortlisted number of open-source LLMs that would be suitable for our customers' use cases. As a result, we have adopted a foundational [Meta Llama model](#) that has been fine tuned for OneAdvanced AI with the OneAdvanced tone of voice for general organisational use.

We have also implemented [Llama Guard 3,](#) which is a Llama-3.1-8B pretrained model, fine-tuned for content safety classification. It acts as an LLM – it generates text in its output that indicates whether a given prompt or response is safe or unsafe, and if unsafe, it also lists the content categories violated.

Llama Guard 3 was aligned to safeguard against the McCommon's standardised hazards taxonomy and designed to support Llama 3.1 capabilities. Specifically, it provides content moderation in 8 languages and was optimised to support safety and security for search and code interpreter tool calls.

In line with our commitment to customers to offer an optimised generative AI solution, we will continue to review the latest evolutions of open-source LLM developments to ensure we are making available to most suitable offering to UK organisations.

Customer Data

Customers are able to privately and securely upload files temporarily within a chat, or permanently within their Private Spaces. When the chat with the temporarily uploaded files is deleted, so are the temporary files.

Private Spaces is a feature of OneAdvanced AI that provides secure storage at both individual and organisational levels. Data saved in a user's Private Space remains accessible only to that user, while organisation-wide data is restricted to authorised users within the same organization on OneAdvanced AI. To maintain a safe environment, all files undergo virus scanning prior to upload. Customers will soon be able to create Private Spaces specifically for teams or groups of people, enabling controlled file storage and access for designated team members.

To ensure the security of the service, virus scanning is performed before a file is uploaded in to OneAdvanced AI.

Fairness and bias mitigation

Fairness and bias mitigation were key factors in our assessment of shortlisted open-source AI models to provide to OneAdvanced customers for OneAdvanced AI. Meta has incorporated advanced techniques for bias mitigation in Meta Llama 3, ensuring that the model generates more balanced and neutral content which is crucial for the foundation of OneAdvanced AI. This includes fine-tuning the model to minimise biased outputs and creating systems to review and address any potential biases with use. Along with bias mitigation, OneAdvanced AI has implemented a range of guardrails and content filters to ensure that our responses are safe, respectful, and accurate.

We have also implemented [Llama Guard 3,](#) which is a Llama-3.1-8B pretrained model, fine-tuned for content safety classification. It acts as an LLM – it generates text in its output that indicates whether a given prompt or response is safe or unsafe, and if unsafe, it also lists the content categories violated. Through our observability layer, we can be flagged guardrail violations to ensure appropriate use of OneAdvanced AI.

Llama Guard 3 was aligned to safeguard against the MLCommons standardised hazards taxonomy and designed to support Llama 3.1 capabilities. Specifically, it provides content moderation in 8 languages and was optimised to support safety and security for search and code interpreter tool calls.

We are deeply committed to developing and using AI responsibly. Guided by EU AI Act, ethical principles and compliance standards, we ensure our AI technologies are built with transparency, fairness, and a focus on societal and environmental benefits. On 12 December 2024, we signed the European Commission's AI Pact, solidifying our commitment to ethical AI practices.

Customer feedback is crucial to ensure we keep improving OneAdvanced for the benefit of all our users. Our users can submit feedback directly and simply within OneAdvanced AI, allowing the team to understand, assess and respond. The option to provide feedback is clearly displayed in the menu within OneAdvanced AI.

OneAdvanced AI monitoring and updates

To monitor OneAdvanced AI, we are leveraging OneAdvanced Platform Observability to monitor system health, proactively identify issues, optimise performance, ensuring that our system is running smoothly, efficiently, and effectively, providing the best possible experience for our users.

OneAdvanced follows a defined Software Development Life Cycle (SDLC) which is based on agile Scrum methodology and incorporates security by design. All changes to the software require a Jira ticket and follow a prescribed workflow, this includes approval by Product, testing by Quality Assurance and peer review of code changes. Automated tests and scheduled security scans run prior to production release.

A known list of 1,000 prompts is used to determine model consistency with previous runs. When system is live, usage is monitored through application logs and system metrics - including the number of blocked requests due to breaching acceptable use guards.

OneAdvanced engage with suitably qualified and accredited third-party service providers for penetration testing across our product portfolio, in addition to testing by OneAdvanced's own security compliance team who are aligned to the OneAdvanced AI SDLC.

OneAdvanced are audited by an accredited independent third-party annually to ensure we remain certified ISO/IEC 27001, 9001 and Cyber Essentials Plus. These frameworks are developed in accordance with the law and industry best practice.

Please see the following to view our ISO Certifications: [Compliance OneAdvanced](#)

Please see the following for Cyber Essentials:
https://registry.blockmarktech.com/certificates/af28917b-85e6-4a77-950a-9e81872d40d7/

Please see the following for Cyber Essentials+:

https://registry.blockmarktech.com/certificates/9eaeed8a-7a88-4537-bc9d-953f280a3270/

To demonstrate our dedication of Responsible AI, we are currently working towards ISO 42001 (Artificial Intelligence Management System), a prestigious certification to ensure the continual improvement of our AI systems.

Explainability and user understanding

The data in OneAdvanced AI to add context to the user's requirement is provided solely by the user for the purpose of their interaction with OneAdvanced AI. Users maintain full control of the data sources used by way of temporary file upload in the chat, or permanent file uploads in Private Spaces.

The user must create a suitable prompt with reference to the data source (file) to determine the response (output). This method ensures humans remain firmly in the loop throughout the chat interaction.

We want all users to make the most of a UK sovereign generative AI solution to help support their tasks for their role. To support our end users within OneAdvanced AI, we have a help section which includes guides such as how to create good prompts.

OneAdvanced AI should be used in accordance with OneAdvanced AI Fair Usage Policy.

Human oversight and contestability

The diverse skill set and experience of the OneAdvanced AI team is attestation to the Responsible AI solution created to serve UK organisations with a sovereign AI solution with privacy, safety, and trust at its core.

The below are (at a minimum) the roles and purpose within the OneAdvanced AI team:

- OneAdvanced AI Product Team - Defines the AI strategy aligned to deliverable objectives that meet the real needs of the customer and market trends. Oversees all Go-To-Market activities.
- AI Engineering Team - Ensuring and delivering a high-quality product aligned to the strategy that is secure, scalable, and performant. Roles include but not limited to, software engineers, architects, data engineers, UI/UX, Quality Assurance, DevOps, Cyber Security, and Integration expertise.
- Our commitment to delivering OneAdvanced AI does not begin and end with the Product and Engineering team. OneAdvanced and OneAdvanced AI have dedicated teams aligned to the strategy who firmly put the customer at the heart of decision making for OneAdvanced AI. Appointed Senior Legal Counsel for Data Privacy, a Data Protection Officer, ESG Officer, Marketing and Services team to onboard and support our OneAdvanced AI customers are completely aligned, as building a great product is only part of the OneAdvanced AI experience.
- We've established an internal AI Steering Committee, a cross-functional group with expertise spanning Legal, Risk & Compliance, Product Development, Engineering, Information Security, and Learning & Development. This team ensures that well-informed, ethical decisions guide our AI systems, alongside upskilling employees with AI education.
- Ultimate accountability for OneAdvanced AI is from the Product and Engineering AI Leadership and the OneAdvanced AI Steering Committee.

Should a user contest an output of OneAdvanced AI, there are feedback mechanisms within OneAdvanced AI that lead straight to the OneAdvanced AI team. Alternatively, we have the contact details to reach out to the team.

Contact and support.

For further information or to request explanations regarding specific decisions made by our AI system, please contact us at ai@oneadvanced.com.

Future of OneAdvanced AI

Although OneAdvanced AI will pay dividends taking advantage of the power of generative AI in your organisation, it's on its journey to unlock the value of Agentic AI. We have invested in a solid, secure, safe, private foundation leveraging the OneAdvanced Platform to provide a generative AI solution for UK organisations. With this foundation and our subject matter experts delivering sector specific software to UK organisations as an established software provider, we have the formula to build pertinent AI Agents that power our customer's world of work, fast.

OneAdvanced AI Agents will be designed to complete sector-specific tasks, freeing up employees to focus on high-value, meaningful work. Each AI Agent will require its own assessment, assembly of subject matter experts and sector brains for the use case. Each Agent or sector specific use case will have its own AI Explainability Statement to ensure transparency and understanding.

oneAdvanced

Powering the world of work

Our business software is the trusted choice for critical sectors, including healthcare, legal services, and education. We keep the world of work moving.

Speak to our expert consultants for personalised advice & recommendations, & get support on the products you are interested in.

Contact us!

📞 +44(0) 330 343 4000   🖥 www.oneadvanced.com   ✉ hello@oneadvanced.com