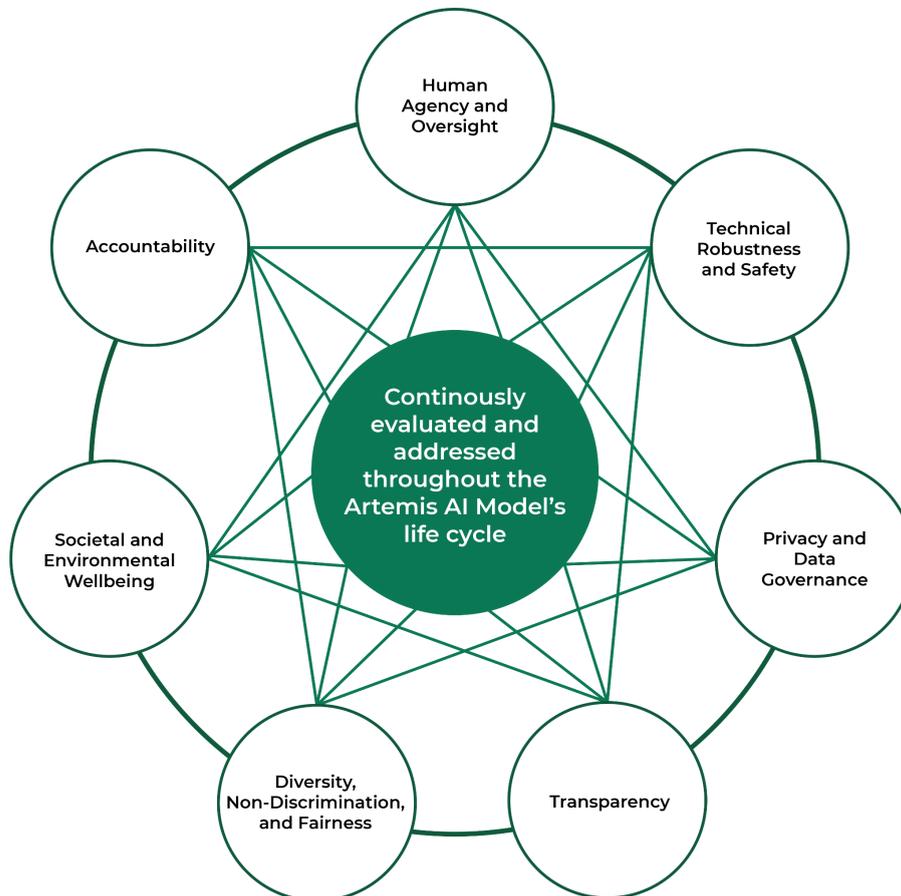


# Responsible AI at Position Green

## SUMMARY



### Position Green's Commitment to Responsible AI

At Position Green, we are committed to developing and using artificial intelligence (AI) in a responsible, transparent, and compliant manner. Our AI systems are designed to support corporate sustainability management and are governed by strong ethical, legal, and technical standards throughout their lifecycle.

In alignment with the EU Artificial Intelligence Act, Position Green's AI systems are classified as limited-risk. We also follow the guidance of European and Swedish regulators, including the European Data Protection Board (EDPB) and Sweden's Integritetsskyddsmyndigheten (IMY). Our policies are regularly reviewed and updated to stay aligned with emerging laws, technical standards, and customer expectations.

**Human Agency and Oversight.** Position Green's AI systems are designed to support, not replace, human decision-making. Our technology is integrated into workflows as an assistive tool, helping users process information more efficiently while keeping all critical decisions in human hands.

We ensure that responsibilities are carefully shared: AI handles routine or data-intensive tasks like classification or analysis, while humans remain in control of interpretation, final decisions, and actions. To avoid overreliance, we provide confidence indicators, source transparency, and clear disclaimers so users can critically assess AI-generated content.

Users can interact with AI at multiple stages, before, during, and after content is generated. They can adjust, verify, or regenerate results to ensure they align with expectations. This promotes accountability and keeps users firmly in control.

Importantly, Position Green's AI does not learn autonomously. All improvements are introduced through structured updates managed by expert teams, ensuring changes are reviewed, validated, and explainable. This maintains a stable, trustworthy, and transparent system, where AI consistently acts as a reliable assistant, never an independent actor.

**Technical Robustness and Safety.** Position Green ensures that all AI systems are developed and maintained with strong safeguards to protect against security threats, misuse, and errors. We follow rigorous security standards for agentive AI features, such as those that can trigger actions or access data, including sandboxing, access controls, input validation, and continuous monitoring. All systems undergo structured threat assessments, with built-in fallback mechanisms and manual override controls to ensure safety and human oversight. Our models are regularly tested for accuracy, reliability, and reproducibility, using both automated and expert-led validation processes. We also monitor model performance in real time and apply structured improvements based on feedback and testing. This layered approach ensures our AI remains secure, accurate, and aligned with Responsible AI principles throughout its lifecycle.

**Privacy and data governance.** Position Green is committed to protecting the privacy and integrity of all data used in its AI systems. We follow a Privacy by Design approach, ensuring data protection is embedded from the outset. Our development practices are guided by SOC 2 controls, aligned with relevant ISO standards, and supported by Data Privacy Impact Assessments. We minimise the use of sensitive data by applying techniques such as pseudonymisation and, where possible, synthetic data. Access to data is tightly controlled and limited to authorised personnel, with full audit trails in place. Users retain full control over their data, including the ability to revoke consent and request deletion of personal or AI-generated data. Our ongoing efforts are focused on enhancing transparency, reducing risk, and aligning with evolving EU privacy regulations and best practices.

**Transparency.** Position Green is committed to ensuring transparency across the entire AI lifecycle, from design and development to deployment and user interaction. We maintain thorough documentation of how our AI systems are built, trained, tested, and validated to ensure full traceability. Our systems are designed with explainability in mind, providing users with source references and clear reasoning where possible, so they can understand and assess AI-generated outputs. All AI features are clearly labelled within our products, and in-product explanations help users interpret results critically. We also provide user-facing resources that clarify how AI components work, their limitations, and when human judgment is essential. User feedback is actively encouraged and used to improve system clarity, performance, and trustworthiness.

**Diversity, Non-Discrimination, and Fairness.** Position Green is committed to building AI systems that are fair, inclusive, and accessible to all users. We actively work to minimise bias in our models by carefully selecting data sources and prioritising equal performance across EU languages, including lower-resource languages. In partnership with Microsoft Azure, we benefit from strong safeguards like content filtering and abuse monitoring to prevent harmful or discriminatory outputs. We also ensure our AI interfaces meet EU

accessibility standards, supporting usability for all individuals. Feedback mechanisms are in place to report any issues related to bias or poor performance, and diverse internal stakeholders are engaged throughout AI development to help ensure inclusive outcomes.

**Societal and Environmental Wellbeing.** Position Green designs its AI systems with sustainability and social responsibility in mind. We minimise environmental impact by selecting low-carbon data centres, partnering with energy-transparent suppliers, and reducing unnecessary use of large language models through efficiency safeguards and content caching. Where possible, we replace resource-intensive models with smaller, more sustainable alternatives. Socially, our AI is designed to be transparent in its interactions and avoids simulating emotional understanding. We also assess broader societal impacts, ensuring our AI supports regulatory compliance, investor transparency, and collective progress on sustainability. By considering both environmental and social effects, we strive to ensure that our AI contributes positively to people, the planet, and public trust.

**Accountability.** Position Green ensures accountability in AI development and use through robust auditability, ethical oversight, and clear reporting channels. All AI processes, including inputs, outputs, and data sources, are logged and traceable, enabling independent audits when required. Our Head of AI oversees compliance with internal guidelines and regular ethics training for developers ensure all stakeholders stay aligned with evolving legal and ethical standards. We also assess trade-offs in AI system design to balance safety, fairness, and utility, supported by Microsoft's content filtering and abuse monitoring systems. Users can report concerns through official channels, and we are committed to investigating and resolving any issues responsibly and transparently.

### **Generative AI: Transparency & Content Integrity**

Position Green complies with the EU's transparency requirements for generative AI. We ensure:

- Users are clearly informed when content is AI-generated.
- AI is designed not to generate illegal or inappropriate content.
- Where applicable, Microsoft (as our model provider) may publish summaries of training data used for general-purpose models.

## 1. INTRODUCTION

Position Green (“we”, “our”, “us”) is committed to the responsible development and deployment of increasingly capable AI systems, including the applications through which users access them. We have put in place, and are continuously investing in, policies and practices to ensure we are appropriately mapping, measuring, and managing AI technology capability and risks across the AI product lifecycle, as further discussed below.

In line with the EU Artificial Intelligence Act, Position Green’s AI systems are classified as limited-risk. We continuously monitor compliance with relevant EU and Swedish regulatory guidance, including recommendations from the European Data Protection Board (EDPB) and Integritetsskyddsmyndigheten (IMY), Sweden’s data protection authority.

References to “Position Green Platform” means an all-inclusive reference to the software developed by Position Green, including: the platform, APIs and any integration or add-on functionalities made available to the clients.

### AI models

Our AI technology is primarily based on Microsoft Azure’s OpenAI models and Azure AI Foundry Models. Additionally, we may make use of Open Source models sourced from the Hugging Face model Hub. We may also develop proprietary models.

#### 1.1. Azure OpenAI models

Position Green uses Azure OpenAI GPT models wherever a generative LLM (Large Language Model) is needed. The models are always deployed within an EU region in the Azure cloud, with minimally the [default content filters](#) enabled. Please refer to Microsoft’s [Transparency Note](#) for more information on how the AI technology works. The models are used as-is, they are not finetuned or adapted to any company data.

#### 1.2. Azure AI Foundry models

Position Green uses Azure AI Foundry models for non-generative AI tasks, such as embedding documents or reranking retrieval results. The models are always deployed within an EU region in the Azure cloud, with minimally the default content filters enabled. Please refer to Microsoft’s [AI Foundry documentation](#) for more information on how the AI technology works. Any generative Azure AI Foundry models are used as-is, they are not finetuned or adapted to any company data. Non-generative models may be finetuned for specific use cases (see 1.4 below for details).

### 1.3. Open source models

The open-source models used at Position Green originally come from the [Hugging Face](#) Model Hub, and are deployed in Position Green's private data center. Please refer to Huggingface's [Content Policy](#) for details concerning content management within the Hugging Face's community. The models may be used as-is or finetuned for specific use cases (see 1.4 below for details).

### 1.4. Proprietary (Non-Generative) AI models

To continuously improve our platform's performance and reliability, Position Green may develop **proprietary, non-generative AI models**. These models are designed to address specialized use cases that cannot be met by existing models available through partners such as OpenAI, Azure Foundry, or Hugging Face.

Our proprietary models are **never generative**, meaning they do not create new text, images, or other content. Instead, they focus on **recognizing patterns, classifying information, or extracting factual data** within a controlled and predefined scope.

These models may be trained or fine-tuned using:

- Publicly available data
- Customer data from organizations that have not opted out of allowing Position Green to use their data for research and development purposes.

The use of such customer data is based on legitimate interest or contractual necessity, in accordance with Articles 6 and 7 of the EU General Data Protection Regulation (GDPR). Customers are informed of this processing and retain the right to object or opt out at any time.

We apply strict safeguards to protect such data:

- Training and testing occur in secure, access-controlled environments.
- Data are pseudonymized or aggregated whenever possible.
- Models are designed so that no customer-specific information can be reproduced or inferred from the output.

Importantly, no personal data is processed by these models, and customer data is never used to train generative systems. This approach ensures that research and development efforts are conducted responsibly, improving model accuracy, robustness, and efficiency, while maintaining full protection of our customers' data and intellectual property.

## 2. TRUSTWORTHY AI ASSESSMENT

In this chapter we have documented all the choices we made that may influence system performance and behaviour, beyond the basics described above.

## 2.1. Human agency and oversight

### 2.1.1. Human agency

**AI in work processes.** At Position Green, we integrate AI into our operations with the aim of augmenting traditional tasks through enhanced decision-making capabilities and expedited navigation of information and documents. The AI system is deployed as an ancillary tool within the workflows, designed to assist human agents by facilitating meaningful engagement and maintaining necessary oversight.

**AI and Human Task Allocation for Enhanced Interaction and Oversight.** In the allocation of responsibilities between the AI system and human personnel, we have deliberately prioritised the assignment of tasks to AI where significant improvements in efficiency can be achieved, such as in data processing and preliminary analyses. Concurrently, we ensure that essential decision-making processes remain under human jurisdiction. This strategic distribution not only capitalises on the respective strengths of human and AI capabilities but also promotes substantive interactions wherein recommendations provided by AI are judiciously reviewed and enacted by human agents.

**Safeguards Against Overreliance on AI in Work Processes.** To safeguard against undue confidence in, and dependence on, the AI system within operational processes, Position Green has instituted specific measures aimed at enhancing transparency. Where possible, confidence scores or uncertainty indicators are included in outputs, as well as human-readable explanations for key outputs. For AI components operating on knowledge bases, the sources from which the AI's responses are derived are presented to the end user. This initiative is intended to encourage users to critically evaluate the AI-provided information, thereby cultivating a deeper comprehension of its operational mechanisms and inherent limitations.

Additionally, users are made aware that they are interacting with an AI component through disclaimers that explicitly remind the user to use the generated output responsibly. This policy not only advocates for critical engagement but also ensures that users maintain an active role in the decision-making continuum, rather than relegating their judgement to the AI system exclusively.

In summary, Position Green's objective is to establish a harmonious and efficacious collaborative dynamic between human employees and AI, where the technology serves to augment human capabilities without supplanting the vital elements of human oversight and critical assessment.

### 2.1.2. Human oversight

**Ensuring Appropriate Human Control in AI Systems.** At Position Green, we place a strong emphasis on maintaining a high level of human control over our AI system, with the aim of boosting user productivity without compromising their independence or ability to make decisions. Our design strategy focuses on ensuring that users remain at the helm when interacting with the AI, granting them the capability to adjust outputs and conduct in-depth source verification. This ensures that the AI functions as a potent tool guided by human oversight, rather than as an autonomous agent making decisions independently.

**Dynamics of Human Involvement.** The concept of "human in control" encompasses all users who interact with the system. Opportunities for human intervention are systematically integrated at various stages of the AI's operation, including:

- › **Before Generating Output:** Users can pose guiding questions to steer the AI's focus (chatbots), or users can select which sources to base the answer generation on (eg. autoreporting).

- › During Interaction: Users have the flexibility to modify the AI's trajectory based on preliminary outputs (chatbots).
- › After Output: Users are encouraged to engage in the analysis and potential refinement of results, leveraging their expertise and specific requirements (all AI functionality).

The degree of human control within our AI system is both significant and foundational to its functionality. Users are empowered to:

- › Modify Output: Directly influence the AI's generated output by requesting adjustments or clarifications, or modifying generated output manually, ensuring outputs align with user expectations.
- › Regenerate Content: Command the AI to produce new responses if the initial output does not meet the user's needs, providing further instructions on the preferred outcome.
- › Source Verification: Engage in a comprehensive review and validation of the sources utilised by the AI, promoting a transparent and trustworthy exchange of information.

Through these mechanisms, our AI system is positioned as a supportive assistant that amplifies human capabilities rather than supplanting human judgement, reinforcing our commitment to a user-centric approach that values and upholds human oversight and expertise.

**Controlled Learning in AI Systems.** In the operational framework of Position Green's Artificial Intelligence (AI) system, our AI does not engage in autonomous self-learning or unsupervised adaptation. The AI's capacity to learn and evolve is strictly governed by controlled updates, rather than through independent self-learning mechanisms triggered by user interactions or feedback on its outputs.

When discrepancies or inaccuracies in the AI's responses are identified, it is important to note that these incidents do not directly contribute to the AI's learning process. Instead, any enhancements or modifications to the AI system are carried out through a deliberate and structured revision process. This process involves a comprehensive review and analysis by our expert team, ensuring that updates to the AI's knowledge base or operational algorithms are meticulously validated and integrated.

This approach is designed to maintain the AI as a stable and consistent tool for our users, providing support based on a fixed repository of knowledge and capabilities, without the unpredictability associated with self-modifying learning systems. Our commitment is to uphold a controlled and transparent environment, wherein the AI serves as a reliable resource, and any advancements in its functionality are introduced in a systematic, supervised manner to guarantee accuracy, reliability, and user trust.

## 2.2. Technical robustness and safety

### 2.2.1. Resilience to attack and security.

Addressing the resilience of our AI system against potential attacks and vulnerabilities is a crucial aspect of our development and operational protocols. As Position Green advances toward more autonomous and agentic AI capabilities, we recognize that our security posture must evolve accordingly. Agentic AI applications (systems that can trigger actions, retrieve data, or interact with internal or external services)

introduce new potential vectors of attack, including prompt injection, data exfiltration, unauthorized access, and model manipulation.

To ensure resilience against such risks, every AI application with agentive or data-accessing capacities must be designed, developed, and deployed in compliance with Position Green's AI Development Guidelines.

**Assessment of Vulnerabilities.** As Position Green develops and deploys AI applications with agentive capacities, we recognize that these systems face a broader range of potential vulnerabilities than traditional, non-agentive AI. Key risks include:

- **Prompt injection and indirect prompt attacks**, through which malicious inputs attempt to override intended instructions;
- **Data exfiltration**, where an AI system could be manipulated to disclose information it has legitimate but restricted access to; and
- **Model manipulation or misuse**, where AI functionality could be repurposed for unintended outcomes.

To mitigate these risks, each AI component undergoes a **structured threat and vulnerability assessment** as part of the AI Development Lifecycle. Assessments are revisited at major release milestones or when functionality changes.

**Integrity and Resilience Measures.** All agentive and data-accessing AI applications are built and operated according to Position Green's AI Development Guidelines. The following core safeguards are mandatory:

- **Isolation and Access Control:** AI agents operate in sandboxed environments under strict, least-privilege permissions. Data sources, APIs, and external services are accessible only through secure, auditable gateways.
- **Input and Output Validation:** By making use of LLMs through Microsoft Azure, Position Green automatically benefits from all security measures put in place by Microsoft, such as prompt shields, content filtering and safety policies, etc. Outputs that may trigger automated actions are checked against whitelists or safety filters before execution. Other AI models used in the Position Green Platform are low-level models that can only execute one specific task, such as identifying the language of a document. Their integration into the AI ecosystem does not introduce any additional potential security weaknesses in the system.
- **Data Protection and Minimization:** Agents access only the minimum data necessary for their function. Sensitive data is encrypted and pseudonymized, and no customer data is retained within the agentive system beyond the execution of the task.
- **Monitoring and Auditability:** All actions performed by agentive systems are logged and continuously monitored for anomalies. Security logs are immutable and reviewed regularly.

**Dual-use Considerations.** Position Green acknowledges the potential for dual-use of AI technologies, where functionalities developed for legitimate purposes could, intentionally or inadvertently, be repurposed for

misuse or harm. To address this, all AI projects undergo a **dual-use review** during design and release stages, considering:

- The scope and autonomy of agentive actions;
- Potential misuse scenarios or access escalation risks;
- Compliance with Position Green's Responsible AI principles and ethical standards.

**Fallback Plan and General Safety.** Human oversight remains an essential safeguard. Every agentive system includes:

- **A safe fallback mode**, allowing human operators to pause or isolate the AI immediately if abnormal behavior is detected.
- **Manual confirmation requirements** for any action that could affect data integrity, client deliverables, or external communications.
- **Incident response procedures**, ensuring rapid containment, impact analysis, and notification in case of a suspected breach or malfunction.

Through layered technical controls, proactive risk assessments, and continuous monitoring, Position Green ensures that all AI systems, including those with limited autonomy, remain secure, controllable, and aligned with the principles of Responsible and Trustworthy AI.

### 2.2.2. Accuracy

**Validation.** A validation process has been established to measure and validate the accuracy of the AI systems. This process is a crucial component of our approach, ensuring that the system's outputs meet the expected standards of reliability. For each use case, a team of subject matter experts designs a test set to evaluate the accuracy of the corresponding AI component. This test set is used for the automated assessment of multiple approaches. Once the optimal approach is identified, subject matter experts conduct a final round of testing and evaluation to confirm the automated evaluation results. If, at any stage of the evaluation process, the accuracy is found to be insufficient, the development strategy is adjusted accordingly to address the issue.

**Accuracy.** To address the inherent imperfections of AI systems, the Position Green Platform incorporates a mechanism to gather indirect and direct feedback on the quality of the generated content. This feedback is invaluable for assessing the frequency and severity of inaccurate predictions, understanding their causes, and addressing them effectively.

Based on insights gained from feedback and other evaluations, we have implemented a systematic approach to continuously enhance the system's accuracy. The steps include:

- Identifying error patterns;
- Formulating hypotheses concerning causes of error patterns;
- Identifying potential solutions, which may involve addressing data quality issues, data shortages, replacing or modifying model architectures, enhancing data processing pipelines, or developing entirely new components;
- Prioritising and testing of potential solutions; and

- Making improvements available to end-users after multiple rounds of automated and human validation.

### 2.2.3. Reliability and reproducibility.

**Monitoring and Testing Strategy.** Position Green has implemented a comprehensive strategy to ensure that every AI component performs reliably, safely, and in accordance with its intended purpose. This strategy combines proactive testing with continuous monitoring to maintain high standards of accuracy and robustness throughout the AI lifecycle.

Testing activities include:

- **Functional Integrity:** Verifying that AI and software components operate correctly and as designed.
- **Resilience:** Testing the system's ability to handle unexpected or atypical inputs without producing erroneous or unsafe behavior.
- **Performance and Scalability:** Assessing computational efficiency, speed, and adaptability to larger datasets or increased workloads.
- **Accuracy and Quality:** Evaluating the precision and reliability of AI-generated outputs across representative use cases.

In addition to structured testing, Position Green operates automated monitoring systems that track the performance of deployed AI models. These systems use both quantitative indicators and qualitative assessments to detect anomalies or potential failures. Any irregularities are reviewed by technical and governance teams to determine causes, assess risks, and inform targeted improvements. This continuous feedback cycle supports ongoing learning and system enhancement.

**Reproducibility in specific contexts.** To ensure consistent and predictable AI behavior, Position Green applies technical measures that minimize randomness and context-dependent variation. Where feasible, model computations use fixed seeds and deterministic settings (for example, zero-temperature generation). These controls promote reproducibility of results across identical inputs and conditions.

**Verification methods.** A structured verification framework is used to evaluate the reliability, reproducibility, and overall performance of AI components. This framework defines measurable criteria for comparison between system versions and supports traceable evaluation across development iterations. Through systematic verification, inconsistencies can be identified early, corrected efficiently, and validated before deployment.

**Communication of failures.** Software bugs and other issues can be communicated to the relevant Position Green team through the contact persons of the customers (customer success managers, account executives, etc.), through the dedicated email-address [support@positiongreen.com](mailto:support@positiongreen.com) or through the [ticket system](#).

## 2.3. Privacy and Data governance

**Respect for privacy and Data Protection.** As AI models become more sophisticated and capable of processing vast amounts of data, at Position Green ensuring privacy and data protection becomes increasingly critical.

- Privacy by Design: Incorporate privacy considerations into AI system development from the outset.
- Data Privacy Impact Assessments (DPIAs): Conduct DPIAs to assess privacy risks and implement appropriate measures.
- Regular Updates: Keep up-to-date with evolving data protection regulations and best practices.
- Ethical AI: Develop and use AI models ethically and responsibly.

**Assessment of Data Types and Scope.** We are committed to rigorously assessing the type and scope of data included in the datasets used for Research and Development of AI features to ensure that no personal or confidential data is contained within them.

**Minimisation of Sensitive Data Usage.** We prioritise the development of AI systems and model training processes that minimise or eliminate the use of potentially sensitive data. By leveraging synthetic data and privacy-preserving techniques such as pseudonymisation, we aim to uphold data privacy while maintaining the efficacy of our AI solutions. Our methodologies are regularly reviewed and updated to ensure the minimal use of sensitive data, reflecting our ongoing commitment to responsible AI development.

**Mechanisms for Notice and Control.** Our policy ensures that users have full control over their data, including mechanisms for obtaining valid consent and the ability to revoke consent when applicable. We provide clear information about data rights, thereby enhancing transparency and trust. Users can contact Position Green through the official support channels and request their data, including all AI-generated data, to be deleted.

**Privacy Enhancement Measures.** While we currently do not employ data masking or obfuscation techniques, we recognise their importance in enhancing data privacy. We are committed to evaluating the feasibility of incorporating these measures into our practices. By investigating and planning for the gradual adoption of these privacy-enhancing techniques, we aim to further strengthen our data protection measures. Staff training on the significance of these methods is also a priority to ensure a comprehensive understanding of their role in safeguarding data privacy.

**Internal Privacy Oversight.** Privacy and data protection are embedded in our AI development practices. Our SOC 2 controls guide how we manage and review personal data processing activities. The Data Protection Officer collaborates with development and compliance teams to ensure privacy by design and ongoing adherence to our internal governance framework.

**Quality and integrity of data.** We are committed to maintaining the quality, integrity, and security of the data used in our AI systems. Our data governance approach is guided by our SOC 2 control framework, with alignment to relevant ISO standards where applicable, particularly in areas of information security, data management, and operational controls.

Oversight mechanisms for data collection, storage, processing, and use are being strengthened through defined procedures, internal audits, and regular control reviews. We also assess the quality and provenance of external data sources to ensure reliability, reduce bias, and maintain trust in our AI outputs.

In collaboration with security and data governance experts, we continue to enhance these processes and align them with established best practices to ensure that our data remains accurate, consistent, and protected throughout its lifecycle.

**Access to data.** Access to data is strictly controlled through role-based (RBAC) and attribute-based (ABAC) access management frameworks, ensuring that only authorized personnel can view or modify data relevant

to their responsibilities. Access is granted on the principle of least privilege and is subject to periodic review by our security and compliance teams.

Any changes to data, including creation, modification, or deletion, are automatically captured in audit logs that record who made the change, when it occurred, and what data was affected. These logs are regularly reviewed to detect and prevent unauthorized activity and to support accountability.

Access to client data is limited to specific roles, such as backend engineers and security specialists, and only permitted for clearly defined purposes (for example, resolving client-reported issues or conducting approved research and development activities). All personnel with such access receive regular training and must adhere to our internal data protection and privacy policies.

Our practices align with SOC 2 controls for security and confidentiality, as well as relevant ISO standards for access control and information security, ensuring robust governance and transparency throughout the data lifecycle.

## 2.4. Transparency

**Traceability.** We acknowledge the importance of establishing measures to ensure traceability within our AI systems. To this end, we aim to document the methods used for designing and developing our algorithmic systems. This includes detailing, where applicable, the programming methods and flow-building processes, the training algorithms, including the selection and gathering of input data.

In terms of testing and validation, we outline the scenarios and cases used to test and validate the AI systems, and document the data used for testing and validation.

By implementing these measures, we strive to enhance the transparency and accountability of our AI systems, ensuring a comprehensive record of their development, testing, and outcomes.

**Explainability.** We conduct assessments to determine how well the decisions and outcomes of our AI system can be understood. In collaboration with a team of sustainability experts, during the testing and evaluation phase of AI components, we meticulously verify the model's outputs to ensure they can be traced back to their inputs and are as expected.

We strive to provide users with explanations for the system's choices by referencing the sources used to generate responses and providing details on the model's reasoning where applicable, ensuring transparency and understanding. From the start, we design the AI system with interpretability in mind, allowing users to track the sources that inform the model's outputs. While the models used by the AI components are not inherently interpretable due to their complex architecture, access to and interpretation of the input data remains possible at all times. This approach helps balance the need for interpretability with the requirement for high accuracy in our AI systems.

**Communication.** All AI-driven features within Position Green's products are clearly identified and labelled as AI. This ensures users are always aware when they are interacting with, or receiving output from, an AI system. Transparency is further supported through in-product explanations and, where relevant, the display of source materials or references used by AI components to generate their results.

Designers and developers are responsible for maintaining clarity, accessibility, and user comprehension in all AI interactions. In addition to communicating what the AI does, they must implement guardrails that help

users interpret results critically, avoiding over-reliance on or blind acceptance of generated outputs. User interface design should actively reduce the risk of confusion, confirmation bias, or cognitive fatigue, for example by providing context, confidence indicators, or reminders to verify important results.

Users are invited to provide feedback directly through integrated feedback channels. All feedback is reviewed and used to guide improvements in system accuracy, reliability, and communication design.

Finally, user guides and help materials consistently highlight the characteristics, limitations, and potential risks of AI features. These resources remind users to apply critical judgment, validate key outputs, and understand the scope and intent of the AI's assistance. Through this approach, Position Green ensures that AI use remains transparent, informed, and responsible.

## 2.5. Diversity, non-discrimination and fairness

**Unfair bias avoidance.** We assessed and acknowledged the possible limitations stemming from the composition of the datasets used by our AI systems. The nature of our AI use cases and relevant data sources, including official legislative documents and company policies, help mitigate the risk of unfair bias. However, a general bias exists in Natural Language Processing (NLP) AI models because they were primarily developed using English-language data. Despite substantial efforts by the AI community to address this, AI systems handling linguistic data remain less efficient and accurate for non-English languages.

For example, interacting with Large Language Models (LLMs) like Azure OpenAI's GPT models is typically more costly for non-English languages. This is because the cost depends on the number of input and output tokens, and non-English inputs convert to more tokens than equivalent English inputs due to historical development choices. To counteract this, Position Green's pricing model is language-neutral and not affected by the user's language or token volume.

Additionally, Position Green explicitly focuses on enhancing the accuracy of its AI components for EU languages other than English, with special attention to low-resource languages, ensuring that non-English users receive the same quality responses as English users.

Users can flag issues related to bias, discrimination or poor performance of the AI system through the contact persons of the customers (customer success managers, account executives, etc.), through the dedicated email-address [support@positiongreen.com](mailto:support@positiongreen.com) or through the [ticket system](#).

**Fairness.** Fairness has been applied in the designing for our AI system through Azure OpenAI. Microsoft's approach to fairness in AI systems includes content filtering and abuse monitoring mechanisms to ensure the responsible use of AI. Their definition of fairness focuses on preventing harmful outputs such as hate speech, violence, and other inappropriate content. This working definition aligns with broader industry standards but is tailored to the specific context of content generated by AI systems, emphasising the prevention of abuse and harm. Microsoft employs quantitative metrics to measure and test fairness, particularly through their content filtering system. This system categorises and assesses the severity of content across multiple dimensions, such as hate speech, sexual content, and violence. Mechanisms established by Microsoft include real-time content filtering and abuse monitoring, which flag and filter inappropriate content. These mechanisms are periodically reviewed and updated to address new challenges and ensure they remain effective.

**Accessibility and universal design.** Our primary user audience comprises corporate entities and

professionals actively engaged in sustainability management. While this focus shapes the design of our AI systems, Position Green strives to ensure that all AI interfaces are accessible and usable for all individuals, in line with EU accessibility standards and best practices.

We recognise the importance of inclusivity and accessibility as fundamental principles of responsible technology development. Therefore, we are committed to the ongoing evaluation and improvement of our AI systems to align with evolving accessibility guidelines and to support a positive and equitable user experience.

**Stakeholder participation.** We have established a mechanism to ensure the active participation of diverse stakeholders in the development and use of our AI systems. This collaborative effort involves close coordination between our AI, development, product, advisory and customer success teams, fostering an inclusive environment for varied perspectives and expertise.

## 2.6. Societal and environmental wellbeing

**Environmental impact.** To address the environmental impact of our AI system, we prioritise selecting data centres with minimal carbon footprint during development and deployment. We aim to use Azure regions known for their lower environmental impact, thus reducing the carbon footprint of our AI system. When evaluating or engaging with new technology suppliers, environmental transparency and sustainability performance are key selection criteria. We prioritise partners that openly disclose their energy efficiency, carbon accounting methods, and commitments to renewable energy use.

In recognition of the high environmental impact associated with the use of Large Language Models (LLMs) such as Azure OpenAI's GPT models, we have implemented a series of mechanisms within our AI components to ensure that these models are only activated when absolutely necessary. Specifically, we perform preliminary checks to verify that all required input data is available and of sufficient quality, and we incorporate caching mechanisms to prevent the regeneration of identical content. Furthermore, we are committed to replacing the use of LLMs with smaller, more environmentally friendly models wherever feasible.

**Social impact.** In addressing the social impact of our AI systems, we prioritise clarity and transparency in its interactions with humans. Our AI systems are designed to clearly signal that its social interactions are simulated and devoid of "understanding" and "feeling." While we have not conducted quantitative assessments regarding attachment or empathy development, we maintain a stance that it does not encourage such connections. Additionally, we recognise the importance of understanding potential social impacts, such as job loss or de-skilling of the workforce.

**Society and democracy.** Recognizing the interconnectedness of stakeholders within the sustainability ecosystem, we have assessed how our AI system may indirectly affect various parties, such as regulators, investors, communities, and other relevant entities. Our AI system contributes to greater transparency and accountability in corporate sustainability practices. This, in turn, can positively influence regulatory compliance, investor decision-making, and community engagement. Moreover, the ability of our AI system to analyse industry peers' sustainability efforts fosters knowledge-sharing and benchmarking, potentially catalysing collective action toward greater sustainability across various sectors. By considering these broader societal impacts, we aim to maximise the contribution of our AI system to advancing sustainable practices while minimising any unintended consequences on affected stakeholders.

## 2.7. Accountability

**Auditability.** To ensure the auditability of our AI system, we have implemented mechanisms that prioritise traceability and logging of the AI system's processes and outcomes. Specifically, we maintain detailed logs of the questions posed, the answers generated, and the sources used throughout the system's operation. We have also ensured that the AI system can be audited independently. Depending on the auditor's requirements, they have access to various components crucial for auditing, including but not limited to the training data, the questions raised by clients, the answers generated, the sources used, the traces, and the underlying models.

**Minimising and reporting negative Impact.** The Head of AI is responsible for overseeing the implementation of our AI Development Guidelines, providing guidance on ethical AI practices, and monitoring Position Green's AI-related activities for compliance with these principles. The Head of AI will play a key role in expanding our training efforts to ensure that all relevant stakeholders are equipped with the knowledge and skills necessary to uphold ethical standards and navigate potential legal frameworks applicable to Position Green's AI components.

All AI developers undergo mandatory ethics training in AI upon onboarding, and to maintain and enhance their understanding, yearly follow-up training is provided. This includes refreshers and updates on new legal frameworks and developments. This ensures the team is well-prepared to address ethical challenges and navigate complex or emerging grey areas.

**Documenting trade-offs.** Azure OpenAI employs several mechanisms to identify relevant interests and values implicated by its AI systems and to navigate potential trade-offs between them. First, Azure OpenAI implements robust content filtering and abuse monitoring systems designed to prevent harmful outputs, such as hate speech and violence. These mechanisms ensure that the AI systems align with ethical standards and societal values, preventing misuse and harm. Secondly, Microsoft uses quantitative metrics to measure and test various aspects of its AI systems. This includes categorising and assessing the severity of content to ensure that the AI operates within acceptable ethical boundaries.

Microsoft further collaborates with experts from various fields to ensure that its AI systems take into account a broad range of societal and ethical implications. This interdisciplinary approach helps them identify and address potential trade-offs between different interests and values, ensuring a holistic perspective in AI system design and deployment.

**Ability to redress.** Users can easily report any potential harm they may have experienced by using the AI functionality on the Position Green Platform through our official support channels. Position Green will acknowledge and assess such reports, and provide information on resolution steps.

## 3. GENERATIVE AI TRANSPARENCY REQUIREMENTS

### 3.1. Disclosing that the content was generated by AI

In adherence to our commitment to transparency and ethical AI usage, Position Green acknowledges the importance of informing individuals when they are interacting with AI-generated content. Unless it is explicitly evident from the circumstances and context of use, Position Green will ensure that individuals are notified of AI involvement. To facilitate this disclosure, Position Green may affix content credentials to the

generated output, allowing individuals to identify that the content was created using AI technology.

### 3.2. Designing the model to prevent it from generating illegal content

We employ prompt engineering techniques to steer the model towards generating responses relevant to Environmental, Social, and Governance (ESG) topics. In the event that a user attempts to prompt the model with queries unrelated to ESG, the system is programmed to recognise this divergence and will respond accordingly by indicating a lack of available sources for such inquiries.

Moreover, our use of GPT models through Microsoft Azure is complemented by its mechanisms designed to deter the generation of inappropriate or illegal content. These built-in safeguards collectively reinforce the model's capacity to produce lawful and ethical content.

### 3.3. Publishing summaries of copyrighted data used for training

Since Position Green uses Azure OpenAI models as provided by Microsoft, without any additional training or fine-tuning, the publication of summaries of copyrighted training data by Position Green is not required. Any public summary of the training content for these general-purpose AI models may be made available by Microsoft. Position Green's responsibility is limited to ensuring the responsible and compliant use of such models, including the prevention of any unauthorised use or reproduction of copyrighted material within AI-generated outputs, and to maintaining full compliance with Microsoft's applicable terms of service.

\*

**Approving body:** CCO, COO, CFO, CMO and CPO

**Approval Date:** 16 Dec. 2025

## APPENDIX

### a. Examples of non-generative AI

- Classification: Assigning a category or label to a document
  - *Document* → *Business Conduct Policy*
  - Language recognition
- Calculating similarity between two inputs
  - *Question vs. paragraph* → *78 % similarity*
- Information Extraction
  - *"Position Green is a Malmö-based company"* → *Position Green = Company Name; Malmö = City*
- Scenario analysis:
  - Estimating ROI of actions
  - Simulating environmental impact of actions