



Standard Bank
Corporate & Investment Banking



Stellenbosch
UNIVERSITY
IYUNESITHI
UNIVERSITEIT

SBSA **RESPONSIBLE** **AI QUESTIONNAIRE**

DISCOVERY PHASE – KEY QUESTIONS

PURPOSE AND NECESSITY

Does the proposed use case genuinely require an AI-based solution?

Not all problems are best solved with AI/digitally; consider cheaper non-AI/digital alternatives given added ethical risks.

Are the intended outcomes aligned with the product's purpose and capabilities?

Check fit vs features, goals, strengths/weaknesses, compatibility, UX and risk—requires a thorough study of both.

Has the business owner validated the value, feasibility, and strategic fit of the product?

Validation implies alignment to objectives plus review of revenue model, scalability, risk and value proposition.

BENEFICENCE (DOING GOOD)

Is the client's problem clearly understood?

Active listening, open questions, empathy, and prioritising pain points by impact improves solution relevance.)

How will the product benefit the client and why is their participation worthwhile?

Use should lead to good outcomes; client contributions (data/feedback/time) should yield meaningful benefit.

How will the product benefit the organisation and other stakeholders?

Organisation also carries risk; benefits should be clear for all relevant stakeholders.

Is there a reasonable balance of risk and reward?

Consider risk identification, risk tolerance, cost–benefit, crisis readiness, and delivery horizon.

Does the product contribute to broader human well-being, including sustainability?

Evaluate beneficence beyond immediate users—societal/environmental impacts such as emissions reduction matter.

NON MALEFICENCE (AVOIDING HARM)

What harms could arise (privacy, bias, unfairness, exploitation, societal/environmental harm)?

Stakeholders include developers, users and those affected by product operation and outcomes; anticipate risks early.

Have intended and unintended consequences been considered?

Balance positives like productivity/cost savings against negatives like employment displacement, opacity, security issues, bias.

Have you peer reviewed ideas and assumptions before proceeding?

Peer review improves and validates thinking; ensure the review group is diverse and appropriately skilled.

Can harms be avoided or mitigated (and are any harms never justified)?

Some harms (e.g., rights violations) can never be justified; others (e.g., unfairness) may be reduced by implementing mitigating measures such as human override.

Given the above, is it ethically advisable to continue?

Some harms are never acceptable; decisions must consider the full stakeholder impact and anticipated outcomes.

BIAS AWARENESS

Have you considered whether bias is affecting your decision-making?

Bias is common; impartial treatment of stakeholders is key; diverse teams can help detect different bias types.

Have you taken steps to reduce bias risk?

No product is bias-free; mitigation can include diverse teams/data, transparency, rigorous preprocessing, bias impact assessment, feedback loops.

JUSTICE AND FAIRNESS

Have you treated stakeholders fairly during discovery and conceptualisation?

Clients sacrifice time/insights; assess whether participation remains worthwhile even if the product fails.

Have you reviewed prior commitments/promises to the client?

Trust requires systematic commitment tracking, clear communication, and proactive issue handling.

Is pilot/discovery client representation fair and inclusive?

Avoid limiting opportunity based on client size/location/revenue; adopt inclusive participation.

Is it fair for clients to contribute if competitors may benefit later?

Embed fairness, openness and responsible data use; integrate ethics into decisions to avoid unfair practices.

Are stakeholder interests preserved in a just and equitable way?

Do deliberate and inclusive stakeholder identification/analysis, transparent communication, equitable decision-making, define dispute mechanisms, provide ongoing updates.

Are there residual benefits for major contributors?

Benefits can include recognition, early access, training, networking, priority input—must be clear and contextual.

Have you compiled a list of stakeholder representatives who can approve development and data use?

Define scope; map stakeholders; review org structures; interview; build a complete approval-capability list.

AUTONOMY, CONSENT, TRANSPARENCY, PRIVACY**Have you enabled fully informed consent for clients to participate?**

Respecting autonomy requires accurate, thorough information so clients can decide meaningfully.

Do clients have opt out mechanisms at any stage?

Use consent documentation, check-ins, privacy controls, clear comms and feedback channels to support autonomy.

Have you protected privacy/personal information and aligned with legal and internal policy requirements?

Privacy protection is ethical and legally required; follow internal processes and good data practices; define collection/use/destruction.

Have you considered data sensitivity classification?

Distinguish public vs private data; link to regulatory compliance and protection requirements.

Are mechanisms in place to retrieve/move/store/destroy data when no longer required?

Implement robust governance, retention policies, movement protocols and audits; comply with internal data governance.

Have you been transparent about the process clients are participating in?

Clients need to understand expectations, risks and process to make truly informed decisions.

Do clients know their role and contributions?

Use scope documents, RACI, comms strategy, feedback channels, progress updates and documented agreements.

Have you clearly explained intended outcomes and benefits?

Effective planning and proactive communication improves alignment and satisfaction.

Have you avoided exaggerated expectations and acknowledged possibility of failure?

Autonomy requires an accurate understanding of risks/limitations and not overselling.

Have you provided guidance on commitment levels, timelines and feedback?

Sets expectations; supports smooth collaboration and successful outcomes.

 EXPLICABILITY, READINESS**Have you put mechanisms in place to document and share lessons learned (including to clients)?**

Shared documentation, formal reports, collaborative reviews and feedback sessions support learning and trust.

Do you have learning collateral for clients to understand the process and ethical foundations?

Collateral supports understanding of development processes and ethical concepts guiding them.

DEVELOPMENT PHASE – KEY QUESTIONS

STRATEGIC ALIGNMENT

Does the product reflect portfolio diversity/new market reach/serving under served clients?

Product choices can affect growth, resilience and society; incorporate diversity and inclusion goals.

Does the product fit the broader business strategy?

Check value proposition, market positioning, scalability, operational efficiency and long term sustainability.

FAIRNESS, BIAS & NON DISCRIMINATION

Is the product choice fair and non discriminatory vs other potential products?

Selection must be defensible; feasibility is valid, “favouring a client” is not.)

Could the product discriminate against clients/stakeholders on unjust grounds?

Discrimination can be unlawful and unethical; evaluate sources of bias across the full lifecycle.

Are processes in place to test/monitor/rectify bias through development, deployment and use?

Bias needs active monitoring and correction at every stage.

Consider creating and peer reviewing bias/discrimination checklist?

Signals transparency, strengthens fairness, and improves credibility.

Is the team composition reasonably diverse?

Diversity increases detection of bias/unfairness and improves decisions and skills.

DATA GOVERNANCE, CONFIDENTIALITY & PRIVACY

Does development protect information as prescribed by law, industry standards and internal policies?

Use privacy-by-design, record keeping, audits, retention/disposal policies, security controls, classification and minimisation.

Have you considered data “Chinese walls”/information barriers/confidentiality undertakings?

Prevent sensitive info flow across divisions; reduce conflicts of interest and protect confidentiality.

Do you have consent/legal entitlement to use personal/confidential information?

Even non personal data may be confidential; consider opt out, minimisation, purpose limitation and granular consent which ensures autonomy, trust and transparency.

HUMAN AGENCY & OVERSIGHT

Have you considered existing business execution processes and human impact before implementation?

A holistic people–process–tech view improves change outcomes and reduces negative impacts.

Have you ensured the product won't undermine autonomy (e.g., manipulation, loss of freedom of choice)?

People must be enabled to make informed decisions; avoid manipulating users for others' aims.

Have you determined the appropriate level of human oversight given risk of harm?

No system is error free; humans must monitor, assess, and override when needed.

Have you put proactive steps in place to detect issues before users encounter them?

Testing, monitoring, piloting, cross functional collaboration, continuous improvement reduce harm.

Can clients/end users engage a human for issues and receive effective feedback?

Multiple feedback channels and prompt support improve human engagement and effective issue resolution.

Are staff responsible for oversight adequately trained?

They must interpret outputs, raise alerts and implement fixes for trends/errors.

Have you developed training materials on proper product use?

Improves product value, client satisfaction and long term product success.

TECHNICAL ROBUSTNESS, SAFETY & ACCURACY

Does the product meet cybersecurity and internal security policy requirements?

Ethical products must be safe and secure; monitor safety/security throughout lifecycle and follow internal governance.

Have you obtained necessary security/governance approvals?

Approvals support compliance, risk mitigation, and trust; this is key to responsible deployment.

Have you secured the product against threats across its lifecycle (including misuse/abuse)?

Plan for malicious actors; safety, security and reliability must be continuously managed

Could low accuracy/availability cause adverse consequences, and do you monitor accuracy?

Accuracy must match risk; monitor false positives/negatives and other performance metrics where relevant.

Have you benchmarked accuracy results against best practice?

Benchmarking goes beyond accuracy to generalisation, fairness and transparency.

Have you considered MLOps processes for ongoing monitoring and timely intervention?

Monitor data quality, drift, anomalies, adversarial signals, costs; document and incorporate feedback.

Can you switch off/withdraw access if inaccurate or misdirected outputs occur?

Incident response, alerting, testing, human in loop, and graceful degradation reduce impact, builds trust and reliability.

Is the data high-quality, current, complete and representative of the deployment environment?

Bias/inaccuracy in data can drive harmful outcomes; test/document integrity and address issues at each stage.

Have you considered impacts of linking multiple data sources?

Linking datasets can improve completeness but can amplify bias, create inconsistencies, and raise privacy issues.

Do you have defined data cleaning procedures (without introducing unintended bias)?

Cleaning is iterative; removing context can introduce bias—clean with product purpose in mind.

Have you considered simpler products before complex ones?

Simpler baseline products are easier to build, explain and maintain; can support incremental progression.

Will you validate for reliability/reproducibility, including in environments different from training?

Test robustness when the test environment differs from training context.)

Have you followed model risk governance requirements where applicable?

Governance approvals support regulatory compliance, risk mitigation, security and trust.

PRIVACY-BY-DESIGN

Did/will you implement privacy-by-design across the lifecycle (incl. end-of-life deletion)?

Privacy protects personally identifiable information/individuals; protect both provided and inferred data; regulate access via governance.

Have you used only relevant/necessary features in development?

Reduces overfitting, multicollinearity, privacy exposure and compute cost; improves interpretability, accuracy, efficiency, and adaptability.

Have you anonymised and secured the data adequately?

Identify PII; de-identify information using appropriate techniques; implement appropriate access controls.

Do data points stay within approved platforms/environments?

Use secure APIs, encryption, access controls, authentication and secure protocols to maintain governance and integrity.

Is there a “right to be forgotten” process (request, removal, destruction)?

Even after consent, data subjects may withdraw consent in certain circumstances; anticipate and operationalise deletion.

Do you keep a repository of previous “right to be forgotten” requests?

Supports compliance, transparency, audit readiness, learning and consistency—while respecting privacy.

Have you embedded opt out options at different levels (platform/product)?

Client-centric UX, seamless onboarding, customer support and regular opt out opportunities support autonomy.)

Is the product aligned with internal data governance requirements and best practice?

Governance protocols and privacy standards guide adequate measures beyond legal minimums.

TRANSPARENCY & EXPLICABILITY

Will users be informed when interacting with AI-generated content or automated decisions?

Autonomy and oversight require awareness of automation and decision-making involvement.

Will you be explicit that the product is not providing advice or acting in an advisory role?

Manages expectations, reduces misuse, and supports transparency.

Have you informed stakeholders about purpose, criteria, implications and limitations of outputs?

Communicate benefits, risks/limitations, data used, and error rates to enable correct use, autonomy and human oversight.

Have you considered playing back outcomes to stakeholders before operationalisation?

Enhances transparency, improves feedback and collaboration, supports expectation management.

Have you considered business processes to supplement known limitations?

Stakeholder communication, cross-functional collaboration, incident response, client education, feedback loops and compliance reduce exposure.

Have you established feedback channels and a way to recalibrate based on feedback?

Transparency requires communication; define metrics, thresholds, feedback repository, feedback analysis, recalibration and (where relevant) automation.

Have you developed different ways to explain outcomes to different users?

Different stakeholders need different explanation depth; counterfactuals and attribute influence can help.

Have you considered visuals and plain language, plus illustrative scenarios for clients?

Supports transparency, different learning styles and clearer understanding of complex products.

Do end-users know what data the product uses and for what purpose?

Supports transparency and expectation management; datasheets can record this.

DIVERSITY, ACCESSIBILITY & OUTCOME FAIRNESS

Is input data diverse and does it represent the target audience?

Training data should reflect the deployment environment and be balanced across subgroups.

Have you peer reviewed/external reviewed data inputs for bias?

Improves credibility and identifies bias earlier.

Have you considered best-practice metrics for measuring data diversity?

E.g., demographic/geographic/temporal/source variation, domain coverage, class balance, quality indicators.

Did you test performance across demographic groups and correct issues?

Historic bias and underrepresentation can degrade performance; allow time and processes for correction.

Have you removed unnecessary sensitive variables and checked for proxy correlations?

Use of personal attributes must be justified and legal; remove proxy variables where needed.

Have you defined what “fair outcomes” look like and tested for it?

Fairness is complex; may require different strategies at input or output level.

Are harms/benefits distributed equitably?

Avoid situations where some bear most harm while others gain most benefit; proportionality relative to contribution matters.

Are stakeholders made aware of potential harms?

Education about harms improves trust and user experience.

Is accessibility considered, including for users with disabilities?

User-centric design should enable successful use across different abilities.

ENVIRONMENTAL & SOCIETAL WELLBEING**Have you evaluated and mitigated environmental impact (energy/carbon)?**

Consider lower computational intensity where possible; manage impact across lifecycle.

Will the product affect work arrangements, and is change management in place?

Avoid degrading work quality; employees are stakeholders needing consultation and training.

Have you informed/consulted impacted business partners and protected their work quality?

Collaboration and feedback support better outcomes.

Have you considered additional people/entities that could be harmed and mitigations?

Products can impact beyond immediate users; anticipate and mitigate downstream harms.

ACCOUNTABILITY & AUDITABILITY**Is the product designed to ensure auditability (training process, data sources, logs, outcomes)?**

Auditability supports accountability even if an audit isn't planned; helps detect and mitigate risk.

Is the product description, intended outcomes and example outputs documented?

Acts as knowledge repository and supports transparency and collaboration.

Is monitoring/logging/feedback data stored for audit (without breaching privacy)?

Supports accountability, traceability, transparency and troubleshooting; privacy must remain protected.

Are detection/reporting mechanisms in place for adverse effects?

Monitoring and reporting channels enable responsible oversight, response and remediation.

Can users request human review for automated outcomes?

Autonomy and oversight require the ability for a human to assess and override decisions or outcomes when necessary.

Are redress mechanisms in place for unfair adverse outcomes?

Justice requires unfair decisions be rectified—via remedy, compensation or correction.

APPLICATION PHASE – KEY QUESTIONS

HONEST MARKETING & EXPECTATION MANAGEMENT

Is the product sold based on what it can actually do?

Fairness and transparency in selling builds trust and avoids over-selling.

Are the terms and conditions of use clearly articulated?

Clear terms of use, including limitations and risks supports explainability and creates a positive user experience.

Have you avoided false promises that could mislead users?

Truthful communication and transparency strengthen brand integrity, reputation and user trust.

Can you meet user expectations and have you clearly addressed limitations?

Transparency about limitations supports explainability, positive user experience, trust and client management.

Have you explained cost and financing implications clearly?

Cost transparency is essential for accessibility, explainability, trust and expectation management.

OPERATIONAL RISK, SECURITY & ROLLOUT

Have you identified operational risk/security threats in product delivery?

Use proactive risk assessments, audits, compliance checks and continuous improvement.

Do you have a risk mitigation strategy for surfacing the product or adding features?

Rollback planning, risk assessment, ethical impact assessment, post-launch monitoring and compliance help mitigate risk; ensures safety, robustness and security.

Have you considered SLAs around availability?

Clear SLAs set expectations and ensures safety and robustness; update as technology evolves.

Have you considered redundancy in maintenance?

Redundancy supports robustness, reliability and uninterrupted service during planned/unplanned disruptions.

SUPPORT, FAIRNESS & INCLUSIVITY IN SERVICE

Is there an identified human-in-the-loop for additional support?

Supports accountability, ensures timely assistance and effective resolution.

Have you considered bias in support channels (language, location)?

Fair and inclusive support systems require linguistic/cultural diversity and equitable access.

Are support queries handled fairly (e.g., first come first served)?

Transparent queueing promotes fairness and efficient resource management.

Have you provided marketing material that is user-friendly for diverse groups?

Inclusive materials support fairness, accessibility and an inclusive product experience.

MONITORING, FEEDBACK & CONTINUOUS IMPROVEMENT

Are there mechanisms to measure client value derived from product usage?

Use metrics like ROI, completion time, satisfaction, cost reduction—ensures continued value delivery.

Are there performance monitoring mechanisms in place?

Drift detection preserves reliability over time, product robustness and security; implement dashboards, drift thresholds, benchmarking and periodic retraining where relevant.

Have you set up feedback channels and follow-up communication after tickets/issues?

Robust communication improves user experience, transparency and explainability.

Have you considered real time recalibration based on feedback?

Continuous monitoring, dynamic strategies, segmentation, tagging and adaptive learning support relevance, accuracy and robustness.

Have you considered continuously updating information based on relevancy?

Keeps product accurate, competitive, relevant and user-centric in dynamic environments; supports relevance, accuracy and robustness.

EQUITABLE REACH, CONTENT STRATEGY & CHANNEL ACCESSIBILITY

Do clients have opt out mechanisms at any stage?

Use consent documentation, check-ins, privacy controls, clear comms and feedback channels to support autonomy.

Are all clients aware of new content/features?

Use segmentation, multi-channel comms, targeted messaging, notifications, webinars and demos whilst respecting client autonomy; supports fairness and transparency.

Does the product/content reach clients through convenient channels—and is feedback collected on that?

Surveys, in app feedback, usage analytics, focus groups and check-ins can confirm channel fit.

Is the product agile enough for future channels?

Modular/scalable architecture, microservices architecture, containerisation technologies, API integrations and user-centric design support agility.

Have you considered content diversity to keep clients engaged?

Strategies include personalisation, multimedia variety, localisation, accessibility features, cultural sensitivity.

Have you considered content volume to avoid spamming?

Respect autonomy, fairness and dignity; use frequency caps, segmentation, opt-in, consistent schedule, relevancy controls.

Have you considered real-time recalibration of content based on client feedback?

Strategies include ongoing monitoring, feedback mechanisms, NLP, content tagging, client segmentation and adaptive learning models.

Have you considered client onboarding/activation/reactivation metrics tied to the “job to be done”?

Track completion, activation, drop-off, retention/churn, satisfaction, support resolution time.

Have you considered data/product availability across diverse groups?

Promotes equity and inclusivity; assess bias, fairness, accessibility, multilingual support.