ANNEX: CROSSWALK COMBINED WITH DESCRIPTION FROM INDIVIDUAL ELEMENTS OF THE AI VERIFY PROCESS CHECKLIST

FUTURE OF PRIVACY FORUM

AI RMF 1.0	AI VERIFY TESTING FRAMEWORK	AI VERIFY PROCESS CHECKLIST
GOVERN 1: Policies, processes, procedures, and practices across the organization related to the mapping, measuring, and managing of AI risks are in place, transparent, and implemented effectively.	**Transparency 1.1.1**	Align with (1) the PDPC's Advisory Guidelines on Key Concepts in the PDPA; (2) Guide to Accountability; and (3) Guide to Data Protection Impact Assessments

	Reproducibility **3.2.1, 3.11.1, 3.14.1**	**[3.2.1]** Verify the quality of data used in the AI system. This may include the following: ● Accuracy in terms of how well the values in the dataset match the true characteristics of the entity described by the dataset ● Completeness in terms of attributes and items e.g., checking for missing values, duplicate records ● Veracity in terms of how credible the data is, including whether the data originated from a reliable source ● How recently the dataset was compiled or updated ● Relevance for the intended purpose ● Integrity in terms of how well extraction and transformation have been performed if multiple datasets are joined ● Usability in terms of how the data are tracked and stored in a consistent, human-readable format ● Providing distribution analysis e.g., feature distributions of input data **[3.11.1]** Record the statistical distribution of input features and output results so that divergence during retraining can be flagged. Monitor input parameters and evaluation metrics for anomalies across retraining runs. **[3.14.1]** Continuous monitoring and periodic validation should be conducted even after models have gone live. This includes: ● Model performance, e.g., monitor feature drift, interference drift, accuracy against ground truth ● Application performance, e.g., latency, throughput, error rates

	Safety 4.1.1, 4.3.1	**[4.1.1]** Complete and submit the Assessment of Materiality to the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) and highlight the risks of the proposed AI solution. Document the justifications for decisions on materiality and the application of relevant governance and controls to demonstrate to regulators and auditors that sufficient responsibility has been taken by humans to address potential risks. **[4.3.1]** Assign a reviewer who is familiar with the downstream use case of an AI model to review the model post-deployment. This process should include model cards / documentation to ensure alignment between intended use cases at modeling and post-deployment.
	Security 5.4.1, 5.5, 5.7	**[5.4.1]** Ensure that the development environment has been secured, including trust access controls **[5.5]** Put in place security measures during the Deployment and Monitoring of AI system development **[5.7]** Put in place security measures for End of Life of AI system

	Robustness 6.1.1, 6.5	**[6.1.1]** • Implement measures to ensure data is up-to-date, complete, and representative of the environment the system will be deployed in • Log training run metadata to do comparison in production, e.g., parameters, and version model to monitor model staleness • Monitor production versus training data characteristics at production stage e.g., statistical distribution, data types, and validation constraints, to detect data and concept drift **[6.5]** Establish a strategy to monitor and mitigate the risk of black box attacks on live AI systems
	Fairness 7.2, 7.7, 7.8	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system **[7.8]** Put in place appropriate mechanisms to ensure fairness in your AI system

	Data Governance 8.3.1	Ensure that assessment has been carried out in accordance with the relevant regulatory requirements and/or industry standards. Mitigation steps have been taken.
	Human Agency and Oversight 10.1.2	Implement a data management system to gather and organize relevant information based on the needs of different user roles (e.g., reviewing models and monitoring live systems)
	Organizational Considerations 12.1	[In development]
GOVERN 2: Accountability structures are in place so that the appropriate teams and individuals are empowered, responsible, and trained for mapping, measuring, and managing AI risks.	**Security 5.1**	Ensure Team Competency

	Data Governance 8.1.1, 8.4	**[8.1.1]** Verify the quality of data used in the AI system. This may include the following: ● Accuracy in terms of how well the values in the dataset match the true characteristics of the entity described by the dataset ● Completeness in terms of attributes and items e.g., checking for missing values, duplicate records ● Veracity in terms of how credible the data is, including whether the data originated from a reliable source ● How recently the dataset was compiled or updated ● Relevance for the intended purpose ● Integrity in terms of how well extraction and transformation have been performed if multiple datasets are joined ● Usability in terms of how the data are tracked and stored in a consistent, human-readable format ● Providing distribution analysis e.g, feature distributions of input data **[8.4]** Ensure team competency in data governance

	Accountability 9.1, 9.3.1	**[9.1]** Establish clear internal governance mechanisms to ensure clear roles and responsibilities for the use of AI by the organization. **[9.3.1]** Implement fine-grained access control that aligns with various roles for users: ● Access to code and data for training AI models ● Access to code and data for deploying AI models ● Access to different execution environments ● Permission to perform various actions (e.g., launch training job, review model, deploy model to server) ● Permission to define access control rules and perform other administrative functions
	Human Agency and Oversight 10.1	Ensure that the various parties involved in using, reviewing and sponsoring the AI system are adequately trained and equipped with the necessary tools and information for proper oversight to: ● Obtain the needed information to conduct inquiries into past decisions made and actions taken throughout the AI lifecycle ● Record information on training and deploying models as part of the workflow process
GOVERN 3: Workforce diversity, equity, inclusion, and accessibility processes are prioritized in the mapping, measuring, and	**Transparency 1.2.4**	Consult end users at the earliest stages of AI system development to communicate how the technology is used and how it will be deployed

managing of AI risks throughout the lifecycle.	**Reproducibility 3.3.1**	Maintain a data provenance record to ascertain the quality of data based on its origin and subsequent transformation. This could include the following: ● Take steps to understand the meaning of and how data was collected ● Document data usage and related concerns ● Ensure any data labeling is done by a representative group of labelers ● Document the procedure for assessing labels for bias ● Trace potential sources of errors ● Update data ● Attribute data to their sources
	Safety 4.4	Assess whether the AI system might fail by considering the input features and predicted outcomes to aid communication with stakeholders.
	Robustness 6.1	Put in place measures to ensure the quality of data used to develop the AI system

	Fairness 7.2, 7.4.2, 7.6, 7.9	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.4.2]** Where feasible, consult the impacted communities on the correct definition of fairness (e.g., representatives of elderly persons or persons with disabilities), values and considerations of those impacted (e.g., individual's preference) **[7.6]** Establish a strategy or a set of procedures to check that the data used in the training of the AI model, is representative of the population who make up the end-users of the AI model **[7.9]** Address the risk of biases due to possible limitations stemming from the composition of the used data sets (lack of diversity, non-representativeness), by applying appropriate adjustments on data samples of minorities
	Accountability 9.1.2	For organizations who are using AI across departments, establish an AI governance committee that comprises representatives from data science, technology, risk and product to facilitate cross-departmental oversight for the lifecycle governance of AI systems.
	Human Agency and Oversight 10.2.3	Implement mechanisms to detect if model input represents an outlier in terms of training data (e.g., return some "data outlier score" with predictions)

	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.
GOVERN 4: Organizational teams are committed to a culture that considers and communicates AI risk.	**Transparency 1.1-1.3**	**[1.1]** Provide the necessary information to end users about the use of their personal data to ensure it is processed in a fair and transparent manner. **[1.2]** Where possible (e.g. not compromising IP, safety or system integrity), identify appropriate junctures in the AI lifecycle to inform end users and/or subjects about the purpose, criteria, limitations, and risks of the decision(s) generated by the AI system in an accessible manner. **[1.3]** Provide information to guide end users on the proper use of the AI system in an accessible manner.

	Safety 4.1- 4.6	[**4.1**] Carry out an assessment of materiality on key stakeholders. [**4.2**] Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. [**4.3**] Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. [**4.4**] Assess whether the AI system might fail by considering the input features and predicted outcomes to aid communication with stakeholders. [**4.5**] Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional") [**4.6**] Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.

	Fairness 7.1 - 7.9	**[7.1]** Assess within-group fairness (also known as individual fairness) **[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.3]** Establish a strategy for the selection of fairness metrics that are aligned with the desired outcomes of the AI system's intended application **[7.4]** Define sensitive features for the organization that are consistent with the legislation and corporate values **[7.5]** Establish a process for identifying and selecting sub-populations between which the AI system should produce fair outcomes **[7.6]** Establish a strategy or a set of procedures to check that the data used in the training of the AI model, is representative of the population who make up the end-users of the AI model **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system **[7.8]** Put in place appropriate mechanisms to ensure fairness in your AI system

		[7.9] Address the risk of biases due to possible limitations stemming from the composition of the used data sets (lack of diversity, non-representativeness), by applying appropriate adjustments on data samples of minorities
	Accountability 9.1.1	Adapt existing structures, communication lines, procedures and rules (e.g., three lines of defense risk management model) or implement new ones.
GOVERN 5: Processes are in place for robust engagement with relevant AI actors.	**Transparency 1.2.4**	Consult end users at the earliest stages of AI system development to communicate how the technology is used and how it will be deployed
	Safety 4.5.4	Close the feedback loop by retraining models with ground truth obtained once models are in production.
	Robustness 6.3	Consider whether the AI system's operation can invalidate the data or assumptions it was trained on e.g., feedback loops, user adaptation, and adversarial attacks
	Data Governance 8.3.1	Ensure that assessment has been carried out in accordance with the relevant regulatory requirements and/or industry standards. Mitigation steps have been taken.

	Fairness 7.2, 7.7, 7.4.2	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.4.2]** Where feasible, consult the impacted communities on the correct definition of fairness (e.g., representatives of elderly persons or persons with disabilities), values and considerations of those impacted (e.g., individual's preference) **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system
	Accountability 9.1.1	Adapt existing structures, communication lines, procedures and rules (e.g., three lines of defense risk management model) or implement new ones.
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.

GOVERN 6: Policies and procedures are in place to address AI risks and benefits arising from third-party software and data and other supply chain issues.	**Transparency 1.1.2**	Publish a privacy policy on your organization's website to share information about the use of personal data in the AI system (e.g., data practices, and decision-making processes). The general disclosure notice could include: ● Disclosure of third-party engagement ● Definition of data ownership and portability ● Depiction of the data flow and identify any leakages ● Identification of standards the company is compliant with as assurance to customers
	Reproducibility 3.13	If using a blackbox model or third party model, assess the vendor's claim on accuracy
	Safety 4.4, 4.5	**[4.4]** Assess whether the AI system might fail by considering the input features and predicted outcomes to aid communication with stakeholders. **[4.5]** Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional")

	Data Governance 8.1 - 8.4	**[8.1]** Put in place measures to ensure data quality over time **[8.2]** Put in place measures to understand the lineage of data, including knowing where the data originally came from, how it was collected, curated, and moved within the organization over time **[8.3]** Ensure data practices comply with relevant regulatory requirements or industry standards **[8.4]** Ensure team competency in data governance
	Accountability 9.5	If you are using third-party "black box" models, assess the suitability and limits of the model for your use case.
MAP 1: Context is established and understood.	**Transparency 1.2**	Where possible (e.g. not compromising IP, safety or system integrity), identify appropriate junctures in the AI lifecycle to inform end users and/or subjects about the purpose, criteria, limitations, and risks of the decision(s) generated by the AI system in an accessible manner.

	Safety 4.1 - 4.6	[**4.1**] Carry out an assessment of materiality on key stakeholders. [**4.2**] Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. [**4.3**] Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. [**4.4**] Assess whether the AI system might fail by considering the input features and predicted outcomes to aid communication with stakeholders. [**4.5**] Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional") [**4.6**] Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.

	Robustness 6.1 - 6.5	**[6.1]** Put in place measures to ensure the quality of data used to develop the AI system **[6.2]** Review factors that may lead to a low level of accuracy of the AI system and assess if it can result in critical, adversarial, or damaging consequences **[6.3]** Consider whether the AI system's operation can invalidate the data or assumptions it was trained on e.g., feedback loops, user adaptation, and adversarial attacks **[6.4]** Put in place a mechanism to evaluate when the AI system has been changed to merit a new review of its technical robustness **[6.5]** Establish a strategy to monitor and mitigate the risk of black box attacks on live AI systems
	Fairness 7.3	Establish a strategy for the selection of fairness metrics that are aligned with the desired outcomes of the AI system's intended application
	Accountability 9.1	Establish clear internal governance mechanisms to ensure clear roles and responsibilities for the use of AI by the organization.
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.

	Organizational Considerations 12.2	[In development]
MAP 2: Categorization of the AI system is performed.	**Transparency 1.1-1.3**	**[1.1]** Provide the necessary information to end users about the use of their personal data to ensure it is processed in a fair and transparent manner. **[1.2]** Where possible (e.g. not compromising IP, safety or system integrity), identify appropriate junctures in the AI lifecycle to inform end users and/or subjects about the purpose, criteria, limitations, and risks of the decision(s) generated by the AI system in an accessible manner. **[1.3]** Provide information to guide end users on the proper use of the AI system in an accessible manner.
	Explainability 2.1	Demonstrate a preference for developing AI models that can explain their decisions or that are interpretable by default

	Reproducibility 3.1 - 3.14	**[3.1]** Put in place methods to record the provenance of the AI model, including the various versions, configurations, data transformations, and underlying source code **[3.2]** Put in place measures to ensure data quality over time **[3.3]** Put in place measures to understand the lineage of data, including knowing where the data originally came from, how it was collected, curated, and moved within the organization over time **[3.4]** Trace the data used by the AI system to make a certain decision(s) or recommendation(s) **[3.5]** Trace the AI model or rules that led to the decision(s) or recommendation(s) of the AI system **[3.6]** Put in place adequate logging practices to record the decision(s) or recommendation(s) of the AI system **[3.7]** Reproduce the training process for every evaluated model (except data) **[3.8]** **A**ssess for repeatability by reviewing if the model produces the same output based on the same input (Note: this is not relevant when it's time to retrain the model) **[3.9]** Define the process for developing models and evaluate the process

		[3.10] Establish a strategy for reproducing the input data used in the training process for every model **[3.11]** Establish a strategy for ensuring that assumptions still hold across subsequent model retraining process on new input data **[3.12]** Reproduce outputs of the AI system **[3.13]** If using a blackbox model or third party model, assess the vendor's claim on accuracy **[3.14]** Establish a strategy to continuously assess the quality of the output(s) of the AI system and ensure that the operating conditions of a live AI system match the thesis under which it was originally developed
	Human Agency & Oversight 10.5	Ensure the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) have considered how the AI system is used to benefit humans in decision-making processes.
	Organizational Considerations 12.3	[In development]

MAP 3: AI capabilities, targeted usage, goals, and expected benefits and costs compared with appropriate benchmarks are understood.	**Transparency 1.1-1.3**	**[1.1]** Provide the necessary information to end users about the use of their personal data to ensure it is processed in a fair and transparent manner. **[1.2]** Where possible (e.g. not compromising IP, safety or system integrity), identify appropriate junctures in the AI lifecycle to inform end users and/or subjects about the purpose, criteria, limitations, and risks of the decision(s) generated by the AI system in an accessible manner. **[1.3]** Provide information to guide end users on the proper use of the AI system in an accessible manner.

	Reproducibility 3.1 - 3.14	**[3.1]** Put in place methods to record the provenance of the AI model, including the various versions, configurations, data transformations, and underlying source code **[3.2]** Put in place measures to ensure data quality over time **[3.3]** Put in place measures to understand the lineage of data, including knowing where the data originally came from, how it was collected, curated, and moved within the organization over time **[3.4]** Trace the data used by the AI system to make a certain decision(s) or recommendation(s) **[3.5]** Trace the AI model or rules that led to the decision(s) or recommendation(s) of the AI system **[3.6]** Put in place adequate logging practices to record the decision(s) or recommendation(s) of the AI system **[3.7]** Reproduce the training process for every evaluated model (except data) **[3.8]** Assess for repeatability by reviewing if the model produces the same output based on the same input (Note: this is not relevant when it's time to retrain the model) **[3.9]** Define the process for developing models and evaluate the process

		[3.10] Establish a strategy for reproducing the input data used in the training process for every model **[3.11]** Establish a strategy for ensuring that assumptions still hold across subsequent model retraining process on new input data **[3.12]** Reproduce outputs of the AI system **[3.13]** If using a blackbox model or third party model, assess the vendor's claim on accuracy **[3.14]** Establish a strategy to continuously assess the quality of the output(s) of the AI system and ensure that the operating conditions of a live AI system match the thesis under which it was originally developed
	Security 5.1	Ensure Team Competency

	Data Governance 8.4	Ensure team competency in data governance
	Accountability 9.1	Establish clear internal governance mechanisms to ensure clear roles and responsibilities for the use of AI by the organization.

	Human Agency and Oversight 10.1-10.5	**[10.1]** Ensure that the various parties involved in using, reviewing and sponsoring the AI system are adequately trained and equipped with the necessary tools and information for proper oversight to: ● Obtain the needed information to conduct inquiries into past decisions made and actions taken throughout the AI lifecycle ● Record information on training and deploying models as part of the workflow process **[10.2]** Ensure specific oversight and control measures to reflect the self-learning or autonomous nature of the AI system **[10.3]** Put in place a review process before AI models are put into production, where key features and properties of the AI model are shared and visualized in a way that is accessible to decision-makers within the organization. **[10.4]** Establish a frequency and process for testing and re-evaluating AI systems. **[10.5]** Ensure the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) have considered how the AI system is used to benefit humans in decision-making processes.
MAP 4: Risks and benefits are mapped for all components of the AI	**Reproducibility 3.13**	If using a blackbox model or third party model, assess the vendor's claim on accuracy

system including third-party software and data.	**Safety 4.1- 4.6**	**[4.1]** Carry out an assessment of materiality on key stakeholders. **[4.2]** Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. **[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. **[4.4]** Assess whether the AI system might fail by considering the input features and predicted outcomes to aid communication with stakeholders. **[4.5]** Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional") **[4.6]** Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.

	Data Governance 8.1 - 8.4	**[8.1]** Put in place measures to ensure data quality over time **[8.2]** Put in place measures to understand the lineage of data, including knowing where the data originally came from, how it was collected, curated, and moved within the organization over time **[8.3]** Ensure data practices comply with relevant regulatory requirements or industry standards **[8.4]** Ensure team competency in data governance
	Accountability 9.5	If you are using third-party "black box" models, assess the suitability and limits of the model for your use case.
MAP 5: Impacts to individuals, groups, communities, organizations, and society are characterized.	**Transparency 1.2.4**	Consult end users at the earliest stages of AI system development to communicate how the technology is used and how it will be deployed.
	Safety 4.1-4.3	**[4.1]** Carry out an assessment of materiality on key stakeholders. **[4.2]** Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. **[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment.

	Fairness 7.4.2	Where feasible, consult the impacted communities on the correct definition of fairness (e.g., representatives of elderly persons or persons with disabilities), values and considerations of those impacted (e.g., individual's preference)
	Human Agency & Oversight 10.4, 10.5	**[10.4]** Establish a frequency and process for testing and re-evaluating AI systems. **[10.5]** Ensure the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) have considered how the AI system is used to benefit humans in decision-making processes.
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.
	Organizational Considerations 12.4	[In development]
MEASURE 1: Appropriate methods and metrics are identified and applied.	**Transparency 1.2.4**	Consult end users at the earliest stages of AI system development to communicate how the technology is used and how it will be deployed.

	Safety 4.1- 4.6	**[4.1]** Carry out an assessment of materiality on key stakeholders. **[4.2]** Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. **[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. **[4.4]** Assess whether the AI system might fail by considering the input features and predicted outcomes to aid communication with stakeholders. **[4.5]** Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional") **[4.6]** Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.
	Fairness 7.3, 7.4	**[7.3]** Establish a strategy for the selection of fairness metrics that are aligned with the desired outcomes of the AI system's intended application **[7.4]** Define sensitive features for the organization that are consistent with the legislation and corporate values
	Accountability 9.4	Establish a strategy for maintaining independent oversight over the development and deployment of AI systems

	Human Agency & Oversight 10.4.1	After models are put into production, put in place mechanisms to review the performance of the models on an ongoing basis, either continuously or at regular intervals. Criteria could be time-based (e.g., every 2 years) or event-based (before the launch of a new AI product, after the introduction of new data, operating context has changed due to external circumstances), or when the AI system has undergone substantial modification.
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.
MEASURE 2: AI systems are evaluated for trustworthy characteristics.	**Transparency 1.2**	Where possible (e.g. not compromising IP, safety or system integrity), identify appropriate junctures in the AI lifecycle to inform end users and/or subjects about the purpose, criteria, limitations, and risks of the decision(s) generated by the AI system in an accessible manner.
	Explainability 2.1	Demonstrate a preference for developing AI models that can explain their decisions or that are interpretable by default

	Reproducibility 3.11, 3.14	Establish a strategy for ensuring that assumptions still hold across subsequent model retraining process on new input data Establish a strategy to continuously assess the quality of the output(s) of the AI system and ensure that the operating conditions of a live AI system match the thesis under which it was originally developed
	Safety 4.2, 4.3, 4.5, 4.6	**[4.2]** Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. **[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. **[4.5]** Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional") **[4.6]** Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.
	Security 5.5	Put in place security measures during the Deployment and Monitoring of AI system development

	Robustness 6.1, 6.2, 6.3, 6.5	**[6.1]** Put in place measures to ensure the quality of data used to develop the AI system **[6.2]** Review factors that may lead to a low level of accuracy of the AI system and assess if it can result in critical, adversarial, or damaging consequences **[6.3]** Consider whether the AI system's operation can invalidate the data or assumptions it was trained on e.g., feedback loops, user adaptation, and adversarial attacks **[6.5]** Establish a strategy to monitor and mitigate the risk of black box attacks on live AI systems

	Fairness 7.2, 7.4.2, 7.6, 7.7, 7.8, 7.9	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.4.2]** Where feasible, consult the impacted communities on the correct definition of fairness (e.g., representatives of elderly persons or persons with disabilities), values and considerations of those impacted (e.g., individual's preference) **[7.6]** Establish a strategy or a set of procedures to check that the data used in the training of the AI model, is representative of the population who make up the end-users of the AI model **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system **[7.8]** Put in place appropriate mechanisms to ensure fairness in your AI system **[7.9]** Address the risk of biases due to possible limitations stemming from the composition of the used data sets (lack of diversity, non-representativeness), by applying appropriate adjustments on data samples of minorities

	Data Governance 8.1-8.4	**[8.1]** Put in place measures to ensure data quality over time **[8.2]** Put in place measures to understand the lineage of data, including knowing where the data originally came from, how it was collected, curated, and moved within the organization over time **[8.3]** Ensure data practices comply with relevant regulatory requirements or industry standards **[8.4]** Ensure team competency in data governance
	Accountability 9.1-9.5	**[9.1]** Establish clear internal governance mechanisms to ensure clear roles and responsibilities for the use of AI by the organization. **[9.2]** Establish the appropriate process or governance-by-design technology to automate or facilitate the AI system's auditability throughout its lifecycle. **[9.3]** Define the policy mechanism for enforcing access rights and permissions for the various roles of users. **[9.4]** Establish a strategy for maintaining independent oversight over the development and deployment of AI systems. **[9.5]** If you are using third-party "black box" models, assess the suitability and limits of the model for your use case.

	Human Agency and Oversight 10.1.2	Implement a data management system to gather and organize relevant information based on the needs of different user roles (e.g., reviewing models and monitoring live systems)
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.
MEASURE 3: Mechanisms for tracking identified AI risks over time are in place.	**Reproducibility 3.11, 3.14**	Establish a strategy for ensuring that assumptions still hold across subsequent model retraining process on new input data Establish a strategy to continuously assess the quality of the output(s) of the AI system and ensure that the operating conditions of a live AI system match the thesis under which it was originally developed
	Safety 4.3, 4.6	**[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. **[4.6]** Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.

	Security 5.3 - 5.7	**[5.3]** Put in place security measures during the Verification and Validation of AI system development **[5.4]** Put in place security measures during the Design and Development of AI system development **[5.5]** Put in place security measures during the Deployment and Monitoring of AI system development **[5.6]** Put in place security measures for the Continual / Online Learning Model **[5.7]** Put in place security measures for End of Life of AI System

	Robustness 6.1 - 6.5	**[6.1]** Put in place measures to ensure the quality of data used to develop the AI system **[6.2]** Review factors that may lead to a low level of accuracy of the AI system and assess if it can result in critical, adversarial, or damaging consequences **[6.3]** Consider whether the AI system's operation can invalidate the data or assumptions it was trained on e.g., feedback loops, user adaptation, and adversarial attacks **[6.4]** Put in place a mechanism to evaluate when the AI system has been changed to merit a new review of its technical robustness **[6.5]** Establish a strategy to monitor and mitigate the risk of black box attacks on live AI systems
	Fairness 7.2, 7.7, 7.8	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system **[7.8]** Put in place appropriate mechanisms to ensure fairness in your AI system

	Accountability 9.1.2, 9.1.3	**[9.1.2]** For organizations who are using AI across departments, establish an AI governance committee that comprises representatives from data science, technology, risk and product to facilitate cross-departmental oversight for the lifecycle governance of AI systems. **[9.1.3]** Enable a process to report on actions or decisions that affect the AI system's outcome, and a corresponding process for the accountable party to respond to the consequences of such an outcome.
	Human Agency & Oversight 10.4.1	After models are put into production, put in place mechanisms to review the performance of the models on an ongoing basis, either continuously or at regular intervals. Criteria could be time-based (e.g., every 2 years) or event-based (before the launch of a new AI product, after the introduction of new data, operating context has changed due to external circumstances), or when the AI system has undergone substantial modification.
	Organizational Considerations 12.5	[In development]
MEASURE 4: Feedback about efficacy of measurement is gathered and assessed.	**Transparency 1.2.4**	Consult end users at the earliest stages of AI system development to communicate how the technology is used and how it will be deployed.

	Accountability 9.1.2, 9.4	**[9.1.2]** For organizations who are using AI across departments, establish an AI governance committee that comprises representatives from data science, technology, risk and product to facilitate cross-departmental oversight for the lifecycle governance of AI systems. **[9.4]** Establish a strategy for maintaining independent oversight over the development and deployment of AI systems.
	Fairness 7.2	Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness
	Accountability 9.1.2	For organizations who are using AI across departments, establish an AI governance committee that comprises representatives from data science, technology, risk and product to facilitate cross-departmental oversight for the lifecycle governance of AI systems.

	Human Agency and Oversight 10.1, 10.3, 10.5	**[10.1]** Ensure that the various parties involved in using, reviewing and sponsoring the AI system are adequately trained and equipped with the necessary tools and information for proper oversight to: ● Obtain the needed information to conduct inquiries into past decisions made and actions taken throughout the AI lifecycle ● Record information on training and deploying models as part of the workflow process **[10.3]** Put in place a review process before AI models are put into production, where key features and properties of the AI model are shared and visualized in a way that is accessible to decision-makers within the organization. **[10.5]** Ensure the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) have considered how the AI system is used to benefit humans in decision-making processes.
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.

MANAGE 1: AI risks based on assessments and other analytical output from the MAP and MEASURE functions are prioritized, responded to, and managed.	**Safety 4.1, 4.2, 4.3, 4.6**	[**4.1**] Carry out an assessment of materiality on key stakeholders. [**4.2**] Assess risks, risk metrics, and risk levels of the AI system in each specific use case, including the dependency of a critical AI system's decisions on its stable and reliable behavior. [**4.3**] Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. [**4.6**] Identify residual risk that cannot be mitigated and assess the organization's tolerance for these risks.

	Fairness 7.1 - 7.9	**[7.1]** Assess within-group fairness (also known as individual fairness) **[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.3]** Establish a strategy for the selection of fairness metrics that are aligned with the desired outcomes of the AI system's intended application **[7.4]** Define sensitive features for the organization that are consistent with the legislation and corporate values **[7.5]** Establish a process for identifying and selecting sub-populations between which the AI system should produce fair outcomes **[7.6]** Establish a strategy or a set of procedures to check that the data used in the training of the AI model, is representative of the population who make up the end-users of the AI model **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system **[7.8]** Put in place appropriate mechanisms to ensure fairness in your AI system

		[7.9] Address the risk of biases due to possible limitations stemming from the composition of the used data sets (lack of diversity, non-representativeness), by applying appropriate adjustments on data samples of minorities
	Accountability 9.1	Establish clear internal governance mechanisms to ensure clear roles and responsibilities for the use of AI by the organization.
	Human Agency & Oversight 10.5	Ensure the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) have considered how the AI system is used to benefit humans in decision-making processes.
MANAGE 2: Strategies to maximize AI benefits and minimize negative impacts are planned, prepared, implemented, documented, and informed by input from relevant AI actors.	**Reproducibility 3.14**	Establish a strategy to continuously assess the quality of the output(s) of the AI system and ensure that the operating conditions of a live AI system match the thesis under which it was originally developed
	Safety 4.3, 4.5	**[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. **[4.5]** Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional")
	Security 5.5	Put in place security measures during the Deployment and Monitoring of AI system development
	Robustness 6.5	Establish a strategy to monitor and mitigate the risk of black box attacks on live AI systems

	Fairness 7.2, 7.7, 7.8	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system **[7.8]** Put in place appropriate mechanisms to ensure fairness in your AI system
	Human Agency and Oversight 10.2	Ensure specific oversight and control measures to reflect the self-learning or autonomous nature of the AI system
	Organizational Considerations 12.6	[In development]
MANAGE 3: AI risks and benefits from third-party entities are managed.	**Transparency 1.1**	Provide the necessary information to end users about the use of their personal data to ensure it is processed in a fair and transparent manner.
	Reproducibility 3.13	If using a blackbox model or third party model, assess the vendor's claim on accuracy

	Data governance 8.1 - 8.3	**[8.1]** Put in place measures to ensure data quality over time **[8.2]** Put in place measures to understand the lineage of data, including knowing where the data originally came from, how it was collected, curated, and moved within the organization over time **[8.3]** Ensure data practices comply with relevant regulatory requirements or industry standards
	Accountability 9.5	If you are using third-party "black box" models, assess the suitability and limits of the model for your use case.
MANAGE 4: Risk treatments, including response and recovery, and communication plans for the identified and measured AI risks are documented and monitored regularly.	**Transparency 1.2.4**	Consult end users at the earliest stages of AI system development to communicate how the technology is used and how it will be deployed
	Reproducibility 3.14	Establish a strategy to continuously assess the quality of the output(s) of the AI system and ensure that the operating conditions of a live AI system match the thesis under which it was originally developed
	Safety 4.3, 4.5	**[4.3]** Put in place a process to continuously assess, measure and monitor risks, including the identification of new risks after deployment. **[4.5]** Plan fault tolerance via, e.g., a duplicated system or another parallel system (AI-based or "conventional")

	Security 5.5, 5.7	**[5.5]** Put in place security measures during the Deployment and Monitoring of AI system development **[5.7]** Put in place security measures for End of Life of AI System
	Robustness 6.1 - 6.5	**[6.1]** Put in place measures to ensure the quality of data used to develop the AI system **[6.2]** Review factors that may lead to a low level of accuracy of the AI system and assess if it can result in critical, adversarial, or damaging consequences **[6.3]** Consider whether the AI system's operation can invalidate the data or assumptions it was trained on e.g., feedback loops, user adaptation, and adversarial attacks **[6.4]** Put in place a mechanism to evaluate when the AI system has been changed to merit a new review of its technical robustness **[6.5]** Establish a strategy to monitor and mitigate the risk of black box attacks on live AI systems

	Fairness 7.2, 7.7	**[7.2]** Put in place processes to test for potential biases during the entire lifecycle of the AI system, so that practitioners can act to mitigate biases based on feedback (e.g., biases due to possible limitations stemming from the composition of the used data sets such as a lack of diversity and non-representativeness **[7.7]** Put in place a mechanism that allows for the flagging of issues related to bias, discrimination, or poor performance of the AI system
	Accountability 9.4	Establish a strategy for maintaining independent oversight over the development and deployment of AI systems.
	Human Agency & Oversight 10.4, 10.5	**[10.4]** Establish a frequency and process for testing and re-evaluating AI systems. **[10.5]** Ensure the appropriate parties who are accountable for the AI system (e.g., AI governance committee, AI system owner, and reviewers) have considered how the AI system is used to benefit humans in decision-making processes.
	Inclusive growth, Societal & Environmental Well-being 11.1	Ensure that the development of the AI system is for the beneficial outcomes for individuals, society and the environment.