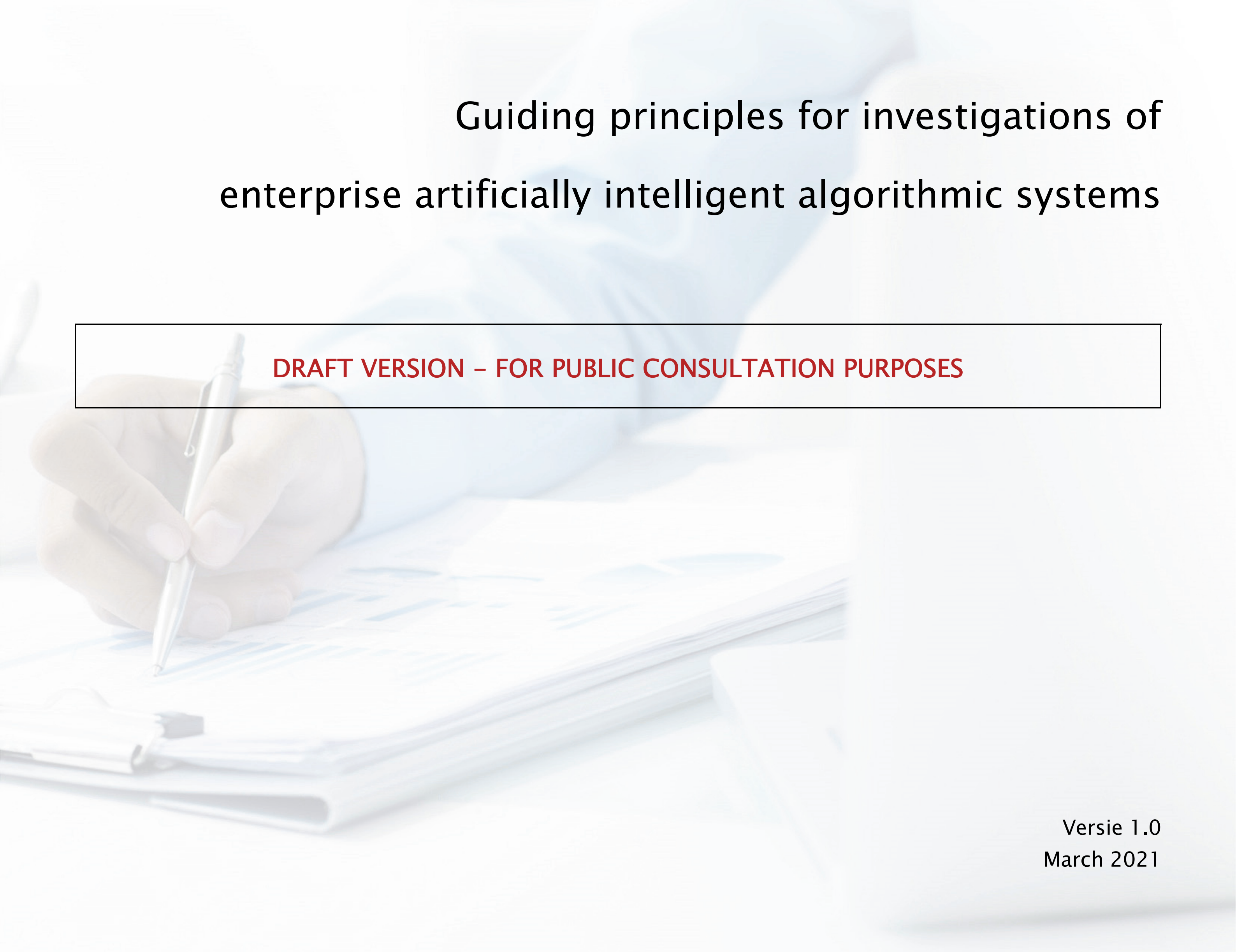
NOREA
DE BEROEPSORGANISATIE VAN IT-AUDITORS

NOREA Guiding Principles Trustworthy AI investigations

Guiding principles for investigations of enterprise artificially intelligent algorithmic systems

DRAFT VERSION – FOR PUBLIC CONSULTATION PURPOSES



Versie 1.0
March 2021

Acknowledgement

This document presents the Guiding Principles for Trustworthy AI investigations (hereafter: “Guiding Principles” and in Dutch: “Studierapport”) developed by NOREA, the Dutch Association of chartered IT-auditors. These principles were developed to guide Dutch chartered IT-auditors (RE’s) in conducting investigations of enterprise artificially intelligent algorithmic systems based on leading practices for Trustworthy AI.

The Guiding Principles were drawn up in English, in order to also be relevant for Dutch organizations operating internationally. A Dutch version will be released once this version is final.

Expert Committee participants

On behalf of the NOREA Expert Committee Algorithm Assurance the following members contributed to the development of these guiding principles:

drs. Mona de Boer RE RA, ir. Tjakko de Boer RE, Damian Borstel RA, ir. Harry van Geijn, mr. Helen Hulsker, Frank van Praat RE MA MSc, drs. Remko Radstake RE RA, Christian Schonenberg RA MSc, Menno Steg RE.

Coordination and editing

Version 1.0: Mona de Boer and Harry van Geijn

PO box 7984, 1008 AD Amsterdam
phone: +31-20-3010380
e-mail: norea@norea.nl
www.norea.nl

Version control		
Version	Date	Amendments
1.0	March 2021	

Table of contents

Section 1 - Introduction

1. Introduction

As part of their digital transformation, organizations increasingly adopt enterprise artificially intelligent algorithmic systems (hereafter: 'algorithmic systems'). For the purpose of these NOREA Guiding Principles, algorithmic systems are defined as computerized mathematical models, used in decision-making processes and characterized by their autonomous inference from vast amounts of varied data sets. The increasing role that these systems have in driving decision making, with significant potential consequences for human welfare (i.e., for employees, consumers, citizens, patients etc.), has given rise to calls for greater accountability in algorithm design, implementation and operation.

This document addresses the role of IT-auditors, as (public) trust providers, in the context of algorithmic accountability. The document presents the Guiding Principles for Trustworthy AI investigations developed by NOREA, the Dutch Association of chartered IT-auditors. These principles were developed to guide Dutch chartered IT-auditors (RE's) in conducting investigations of algorithmic systems based on leading practices for Trustworthy AI.

Both enterprise adoption of algorithmic systems, as well as the scrutiny thereof by IT-auditors are emerging fields. Conscious of this early stage, the Guiding Principles do not (yet) cover the provision of assurance on design, implementation or operational effectiveness of algorithmic systems, as the dialogue on the required algorithmic accountability norms as well as the corresponding audit procedures to provide assurance is still in motion. Instead, the Guiding Principles aim to support IT-auditors in performing ex-ante and/or ex-post investigations of algorithmic systems to guide and support organizations in their journey towards deploying trustworthy AI applications. Furthermore, the principles do not intend to present a comprehensive framework for algorithmic system scrutiny, but are focused on highlighting key questions for the IT-auditor to raise as 'context-relevant', based on current leading practices for Trustworthy AI.

Going forward, the Guiding Principles will be updated by the NOREA Expert Committee Algorithm Assurance, when significant developments in theory as well as in practice occur.

2. Structure of the Guiding Principles

The Guiding Principles are structured according to the Cross Industry Standard Process for Data Mining (CRISP-DM). CRISP-DM is well-known within the data analytics and -science community, resonates, and covers all relevant steps in the development cycle of algorithmic systems. Furthermore, it enables IT-auditors to focus their investigation on specific parts of the development cycle as relevant for organizations and their stakeholders.

For each phase of the CRISP-DM process, risk categories were established that are relevant to the correct functioning of algorithmic systems: Governance, Ethics, Privacy, Performance and Security. Various public, private, and civil organizations have produced frameworks for these risk categories. The NOREA Expert Committee Algorithm Assurance has reviewed a range of these frameworks with the objective to select a comprehensive and accepted industry standard in the European context. Each risk category was detailed into a number of key questions for the IT-auditor to raise, related to control objectives and measures. These questions were derived from the following laws, regulations and leading practices:

1. ISO 24028[1] and 27001[2]
2. European Commission AI High Level Expert Group: Ethics Guidelines for Trustworthy AI[3].
3. GDPR.[4]
4. Information Commissioner's Office Guidance on the AI Auditing Framework[5].
5. COBIT[6].

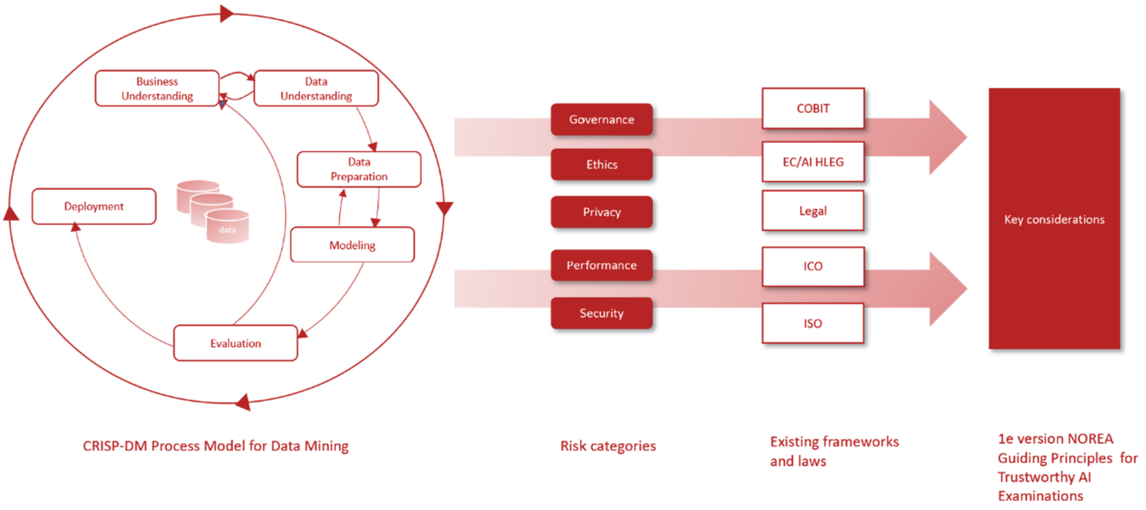


Figure 1: structure of NOREA Guiding Principles for Trustworthy AI investigations

The Guiding Principles and a descriptive numerical summary thereof are included in Section 2 of this document.

[1] ISO/IEC TR 24028, Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence, first edition 2020-05.

[2] ISO/IEC 27001:2013, Information technology — Security techniques — Information security management systems — Requirements, Second edition 2013-10-01.

[3] EC/AI High Level Expert Group on artificial Intelligence, 8 april 2019.

[4] General Data Protection Regulation (EU) 2016/679. Given that this document is in English we refer to GDPR instead of 'AVG', which is the Dutch name for the same regulation. The GDPR is the data privacy law that applies in the EU. When we refer to the GDPR, we do want to point out that there are limited issues subject to national law and therefore some deviations from the GDPR might apply in the Netherlands.

[5] Information Commissioner's Office, draft guidance for consultation, 20200214, version 1.0.

[6] COBIT® 2019 Framework: Governance and Management Objectives.

3. Considerations in using the Guiding Principles

In performing investigations of algorithmic systems and applying these Guiding Principles, there are a number of considerations for IT-auditors to take into account:

- The Guiding Principles also oversee outsourced (sub)processes relevant to the design, implementation and operation of algorithmic systems.

- Both the development and the scrutiny of algorithmic systems are highly multi-disciplinary by nature. This means that it is very likely that IT-auditors will (have to) involve subject matters experts (e.g. data scientists, business ethicists, legal/privacy experts, cybersecurity experts) in conducting their investigations.

With respect to the matters above, we refer to the existing NOREA standards regarding (the use of) service organizations and (the use of) experts.

4. The Guiding Principles and related publications

Dependent on the nature, scope and extent of the investigation of an algorithmic system, IT-auditors could, in addition to these Guiding Principles, benefit from other sources. In this context, the Expert Committee Algorithm Assurance highlights the following publications:

- *NOREA Guide Privacy Control Framework (NOREA, version 2.0, August 2019)*
https://www.norea.nl/download/?id=6038

- *The Machine Learning Audit—CRISP-DM Framework (ISACA, 6 January 2018)*
https://www.isaca.org/resources/isaca-journal/issues/2018/volume-1/the-machine-learning-auditcrisp-dm-framework

- *Global Perspectives and Insights: The IIA's Artificial Intelligence Auditing Framework Part I-III (October 2017 / February 2018)*
https://www.iia.nl/actualiteit/nieuwsglobal-perspectives-and-insights-artificial-intelligence
https://www.iia.nl/actualiteit/nieuws/global-perspectives-and-insights--the-iias-artificial-intelligence-auditing-framework-part-ii
https://www.iia.nl/actualiteit/nieuws/global-perspectives-and-insights-the-iias-artificial-intelligence-auditing-framework-part-iii

5. How the Guiding Principles were established

The Guiding Principles (current version 1.0 for public consultation purposes) were developed by the Expert Committee Algorithm Assurance of NOREA in 2020-2021. This document was peer-reviewed and subsequently approved by NOREA's Professional Practices Committee ("Vaktechnische Commissie") and Board in March 2021 for publication for public consultation purposes. After the public consultation period, changes will be made to the document by the Expert Committee as relevant, and the Guiding Principles will be released in final.

Section 2. Guiding Principles

1. Numerical summary of Guiding Principles

The table below summarizes the Guiding Principles. The Principles contain 119 key considerations for Trustworthy AI investigations, categorized into 5 risk categories and 6 CRISP-DM phases + an added Governance phase. Table 1 presents a descriptive numerical summary of the Guiding Principles, followed by the detailed framework. After the public consultation phase of the Guiding Principles, an interactive version of the detailed framework will be released allowing other cross sections than the current CRISP-DM + Governance phases.

CRISP-DM Phase	Risk categories	# Key considerations
1. Business Understanding	Governance	3
	Ethics	16
	Privacy	4
	Performance	4
	Security	1
2. Data Understanding	Ethics	5
	Privacy	4
	Performance	6
	Security	2
3. Data Preparation	Ethics	1
	Privacy	4
	Performance	1
	Security	5
4. Modeling	Ethics	4
	Privacy	2
	Performance	8
	Security	6
5. Evaluation	Ethics	2
	Performance	22

CRISP-DM Phase	Risk categories	# Key considerations
	Security	1
6. Deployment	Ethics	1
	Performance	8
	Security	2
Added Phase		
Governance	Roles & responsibilities	2
	Ethics	1
	Privacy	2
	Performance	2
Total		119

Table 1: numerical summary of NOREA Guiding Principles

2. Guiding Principles

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
Business Understanding				
1	Governance	ICO	Has the organization defined and documented common language for the development, implementation and operation of its algorithmic systems?	• Investigate if documentation of different types of algorithmic systems and language/"jargon" is in place to describe processes, internal functionality, descriptions, interpretation of results etc.; • Check whether a glossary or dictionary is in place within the organization that all stakeholders involved with algorithmic systems adhere to.
2	Governance	COBIT	Has the organization performed an impact assessment to determine whether the usage of the algorithmic system may negatively impact the existing governance mechanisms over its data processing?	Determine whether a governance framework is in place and used to monitor the data processing of algorithmic systems within the organization.
3	Governance	COBIT	Has the organization performed an assessment to verify whether its algorithmic system is in line with its organizational risk appetite?	Check whether regular reviews are performed regarding how algorithmic systems are used within the organization, how they perform, and whether this is in line with the organizational risk appetite.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
4	Ethics	EC/AI HLEG April 2019 – chapter I.1	Has the organization performed an impact assessment for its algorithmic system to assess potential negative impacts of its operation on fundamental human rights?	• Determine if an impact assessment is performed and documented, including any decisions on potential trade-offs made between the different ethical principles and human rights that were identified. • Determine if human rights like human dignity, freedom of the individual, respect for democracy, justice and the rule of law, equality and citizen's rights are considered.
5	Ethics	EC/AI HLEG April 2019 – chapter II. 1.6	• Has the organization performed an impact assessment for its algorithmic system to assess its broader societal impact (e.g., impact beyond the individual (end)user, such as potentially indirectly affected stakeholders)? • What steps has the organization taken to counteract such risks?	An example can be whether the organization assessed whether there is a risk of job loss or de-skilling of the workforce.
6	Ethics	EC/AI HLEG April 2019 – chapter I. 1.1	If the algorithmic system is used in a work and labor process, has the organization considered the task allocation between the system and humans for meaningful interactions and appropriate human oversight and control?	For example, does the algorithmic system enhance or augment human capabilities? Investigate safeguards the organization has taken to prevent overconfidence in or overreliance on the algorithmic system for work processes.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
7	Ethics	EC/AI HLEG April 2019 – chapter I. 1.1	Has the organization implemented an appropriate level of human control to control the risks of the algorithmic system and its outcomes?	Investigate the documentation of the level of human control or involvement; for example, who is the “human in control” and what are the moments or tools for human intervention. The documentation can also include what the mechanisms and measures are to ensure human control or oversight, and did you take any measures to enable audit and to remedy issues related to governing AI autonomy?
8	Ethics	EC/AI HLEG April 2019 – chapter II. 1.2	If there is a probable chance that the algorithmic system may cause damage or harm to users or third parties, has the organization assessed the likelihood, potential damage, impacted audience and severity?	Verify if the assessment includes: • Liability and consumer protection rules; • The potential impact or safety risk to the environment or to animals.
9	Ethics	EC/AI HLEG April 2019 – chapter II. 1.2	Has the organization assessed what level and definition of accuracy would be required for the algorithmic system in its context?	Investigate how accuracy is measured and assured. This should include measures to ensure that the data used is comprehensive and up to date, and whether there is a need for additional data, for example to improve accuracy or to eliminate bias. It should also include a verification of what harm would be caused if the algorithmic system makes inaccurate predictions.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
10	Ethics	EC/AI HLEG April 2019 – chapter II. 1.4	Has the organization implemented measures to ensure traceability of the algorithmic system?	Examples of traceability are documentation of the following methods: • Methods used for designing and developing the algorithmic system: ○ Rule-based AI-systems: the method of programming or how the model was built; ○ Learning-based algorithmic system; the method of training the algorithm, including which input data was gathered and selected, and how this occurred. • Methods used to test and validate the algorithmic system: ○ Rule-based algorithmic system; the scenarios or cases used in order to test and validate; ○ Learning-based algorithmic system: information about the data used to test and validate. • Outcomes of the algorithmic system: ○ The outcomes of or decisions taken by the algorithm, as well as potential other decisions that would result from different cases (for example, for other subgroups of users).

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
11	Ethics	EC/AI HLEG April 2019 – chapter II. 1.4	Has the organization assessed: • To what extent the decisions and hence the outcome made by the algorithmic system can be explained? • To what degree the system's decision influences the organization's decision-making processes? • Why the particular system was deployed in the specific area? • The system's business case and added value?	Investigate the interpretability after the model's training and development. Assess whether the organization can analyze training and testing data, and if this can be changed and updated over time.
12	Ethics	EC/AI HLEG April 2019 – chapter II. 1.4	• Has the organization implemented mechanisms to inform (end)users on the reasons and criteria behind the algorithmic system's outcomes and who or what may benefit from the product/service? • Did the organization clearly communicate the system's characteristics, limitations and potential shortcomings?	Investigate whether the communication is clearly and intelligibly to the intended audience, whether processes are established that consider users' feedback and use this to adapt the system. An example for about the content, is to verify whether the communication is clear about (i) usage scenarios for the product and clearly communicate these to ensure that it is understandable and appropriate for the intended audience(ii) potential or perceived risks, such as bias.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
13	Ethics	EC/AI HLEG April 2019 – chapter II. 1.5	What measures have been implemented by the organization to avoid creating or reinforcing unfair bias in the algorithmic system, both regarding the use of input data as well as for the system's design?	Investigate whether the review and documentation of the strategy entails (i) possible limitations stemming from the composition of the used data sets, (ii) diversity and representativeness of users in the data, (iii) any test for specific populations or problematic use cases (iv) the research and use of available technical tools to improve your understanding of the data, model and performance (v) a process to test and monitor for potential biases during the development, deployment and use phase of the algorithmic system.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
14	Ethics	EC/AI HLEG April 2019 – chapter II. 1.5	• Has the organization formulated an adequate working definition of "fairness" that is applied during the development and implementation of the algorithmic system? • Has the organization: ○ Considered other definitions before choosing this one? ○ Performed a quantitative analysis to identify metrics to measure and test the applied definition of fairness? ○ Established mechanisms to ensure fairness in the algorithmic system and its outcomes?	Verify if the requirements for fairness are applied in the algorithmic systems, and whether a data scientist to measure and test against the set requirements has been involved and legal advice has been obtained for the formulation of fairness.
15	Ethics	EC/AI HLEG April 2019 – chapter II. 1.5	Is the algorithmic system accessible and usable by a wide range of individuals with various preferences and abilities?	Verify that the design of the algorithmic system is: • Accessible also to users of assistive technologies; • Usable by those with special needs or disabilities or those at risk of exclusion; • Build by a team that is representative of you're the target user audience or if feedback was obtained from other teams or groups that represent different backgrounds and experiences.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
16	Ethics	EC/AI HLEG April 2019 – chapter II. 1.5	Has the organization involved stakeholders that are impacted by the algorithmic system during development and implementation?	Verify whether the organization has informed and involved workers and their representatives impacted by the algorithmic system in advance. Additionally, verify if the developers have completed training so they can identify and address bias and discrimination in AI models.
17	Ethics	EC/AI HLEG April 2019 – chapter II. 1.7	Has the organization established mechanisms to identify relevant interests and values implicated by the algorithmic system and potential trade-offs between them?	Investigate the documentation of the mechanism and process for decisions on such trade-offs.
18	Ethics	general legal	Is the algorithmic system and its underlying technology permitted by law?	Verify if a legal assessment on the legitimacy of the system and technology has been done. Examples of relevant legislation can be single-subject rules on prohibition on the use of facial recognition in specific matters, profiling by the government or even rules on the use of AI for a specific industry.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
19	Ethics	general legal	Has the organization assessed the legal qualification of the algorithmic system, and are the legal consequences of such qualification taken into consideration during development and implementation?	Verify if a legal assessment has been done on applicable law and impact. Relevant legislation can be of general product safety rules and more sector-specific rules covering for example cars, machines, planes and medical devices. Product liability law is often complemented by general liability laws. The legal qualification (e.g., as a product, as rent, as consumer product etc.) should also be considered.
20	Privacy	GDPR, chapter 2 and 4	Has the organization assessed the following privacy implications for the algorithmic system: • Has the purpose (including secondary purposes) of the algorithmic system been identified? • Have all direct and indirect types of personal data of the algorithmic system been identified? • Is there a lawful basis for all the purposes of the algorithmic system and the use of personal data?	Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the assessment of purpose, personal data categories, and the lawful basis for processing.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
21	Privacy	GDPR, chapter 4	Has the algorithmic system been designed to accommodate for the right to object to automated processing (e.g., are there alternative processing possibilities)?	• Investigate (by inquiry of the system owner or inspection of system documentation) whether the algorithmic system is designed to accommodate that data of data subjects is excluded from processing and/or that data subjects are not subjected to automated decision making; • Obtain or generate test documentation and test results to verify the implementation of the right to object functionality.
22	Privacy	GDPR, chapter 2 and 4	• Is the personal data collected proportional, relevant and necessary for the purpose of processing? • Have alternatives been considered using fewer personal data to achieve the same objectives of processing?	Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the proportionality, relevance and necessity of the collected data for the purpose of processing (data minimization).

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
23	Privacy	GDPR, chapter 2 and 4	• Does the algorithmic system produce or support decisions with legal or other significant effects to data subjects? • Has the algorithmic system been designed to accommodate for the right not to be subjected to solely automated decision making? • Has the algorithmic system been designed to be transparent about the basis for decisions/conclusions?	• Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the assessment of whether the type of processing may result in a high risk to the rights and freedoms of data subjects; • Investigate (by inquiry of the system owner or inspection of system documentation) whether the algorithmic system is designed to accommodate that data of data subjects is excluded from processing and/or that data subjects are not subjected to automated decision making; and that transparency of decision support results is embedded in the design of the algorithmic system; • Obtain or generate test documentation and test results to verify the implementation of the right to object functionality and of the transparency functionality.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
24	Performance		• What is the decision the algorithmic system is designed to support? • Was the algorithmic system designed specifically to support this decision, or is an existing algorithmic system being re-used? If so, is this assessed as appropriate?	• Identify who the stakeholders of the decision are; • Identify considerations given to alternative solutions to support the decision.
25	Performance		Is there evidence of the rationale and the scoping of the algorithmic system concept?	Documentation detailing the rationale, concept and structure of the model, such as: • What the algorithmic system aims to replicate; • The input, output and algorithmic system logic; • The algorithmic system type (including options for alternative approaches which have been rejected); • The stakeholders responsible for policy and delivery; • The required precision (offset against complexity); • Identification of the limitations of the model.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
26	Performance	ISO24028	Is there a clear understanding of the requirement for creating, using and/or maintaining the algorithmic system in terms of staffing, resources and skills required?	Defined requirements exist for the training, staffing and skill requirements of human resources involved with usage of the algorithmic system and the interpretation of its results.
27	Performance	ISO24028	Is there a clear understanding of the dependency on the algorithmic system and if applicable its vendor/developer to prevent lock in and dependency on external providers to maintain or use the algorithmic system?	Check for the existence of an exit or change strategy within the design of the development plan for the algorithmic system(s) that takes into account independency from external providers or ways to mitigate vendor lock-in risk.
28	Security	ISO27001 A.5.1	Have AI security risks, attacks and threats been addressed in the current security policies?	• Check whether AI specific risks are addressed in documented security policy such as Poisoning attacks (including mitigating and managing these risk); • Check whether necessary information in case of a risk for human physical integrity is considered. An insurance policy to deal with potential damage from the algorithmic system could be considered.
Data Understanding				

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
29	Ethics	EC/AI HLEG April 2019 – chapter II. 1.3	Has the organization involved the Data Privacy Officer (DPO)? Has the DPO assessed the type and scope of data in training and operation data sets (for example whether they contain personal data)?	Verify if the Data Protection Officer is involved and a Privacy Impact Assessment ("PIA") is initiated. In case personal data is used, ensure that measures to enhance privacy, such as via encryption, anonymization and aggregation are taken and included in the PIA.
30	Ethics	EC/AI HLEG April 2019 – chapter II. 1.3	If the algorithmic system's processes of data collection (for training and operation) and data processing, has the organization established a mechanism allowing others to flag issues related to privacy or data protection?	In case personal data is used for development and/or deployment, it needs to be verified whether the legal grounds for the respective processing are taken into account, and whether additional action is required (e.g., obtain consent). If personal data is used, then it should also be verified whether mechanisms for notice and control over personal data to be built in (such as valid consent and possibility to revoke).
31	Ethics	EC/AI HLEG April 2019 chapter II 1.7	• Has the organization facilitated the algorithmic system to be auditable? • In case of applications affecting fundamental rights (including safety-critical applications), can that be audited independently?	Investigate what the mechanisms are that ensure traceability and logging of the algorithmic system's processes and outcomes.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
32	Ethics	EC/AI HLEG April 2019 chapter II 1.7	Does the organization provide training and education to help develop accountability practices?	The organization can consider training for relevant workers involved, not only for the development phase. Training should include the potential legal framework, if applicable to the AI-system. The organization should also consider establishing an 'ethical AI review board' or a similar mechanism to discuss overall accountability and ethics practices, including potentially unclear grey areas.
33	Ethics	general legal	Has the organization assessed whether unlawful bias can occur in the algorithmic system (input and output)? For the assessment the relevant grounds of discrimination need to be taken into account (e.g., race, nationality, sexual preferences etc.).	There are various laws and regulations on equal treatment that can be relevant for the assessment: • The fundamental right of equal treatment and prohibition of discrimination; • Equal treatment of employees; and • Any sector specific rules. These laws and regulations are often not specific to AI yet, but will apply in the meantime.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
34	Privacy	GDPR, chapter 4	Has user management been implemented on data and the algorithmic system?	Investigate (by inquiry of the system owner or inspection of system documentation) whether the algorithmic system is designed with appropriate user management functionality.
35	Privacy	GDPR, chapter 2 and 4	Does the algorithmic system process any special categories of personal data: racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data, data concerning health, or data concerning a natural person's sex life or sexual orientation?	Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the assessment of special categories of personal data.
36	Privacy	GDPR, chapter 2 and 4	Can the personal data processed be used to profile or discriminate data subjects (like address)?	Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the assessment of whether the type of processing combined with the categories of personal data may result in a high risk to the rights and freedoms of data subjects such as profiling or discrimination.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
37	Privacy	ICO	Is the data collected for the development, training and implementation of the algorithmic system limited to the scope of the solution so as to prevent noisy data and collecting excessive data (which might contains unneeded personal information or restricted data)?	A scoping document exists that clearly defines the requirements of data used for development, training and implementation of the algorithmic system. This should include sources and outline all fields and variables included and needed within the selected data.
38	Performance		Is there a technical guide that demonstrates the logical flow of the algorithmic system?	• Compare the data flow, logic and structure in the model with the description in the technical guide; • Conclude on the understanding and logics of the algorithmic system and discuss remaining questions with experts.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
39	Performance		Is the data in the algorithmic system of good quality?	• Identify if relevant data quality checks are in place and assess the effectiveness; • Review the quality of data and sources, such as the extent to which data: o are up-to-date; o sources are documented; o is based on a robust sample; o is consistent with other sources; and o meets the requirements it is being used for; • Check if data (as much as is practically feasible) in the model align with the source data to conclude on accuracy; • Does model documentation outline the limitations of the used data? • Where good data quality is lacking, determine what steps have been taken to work around this, for example making use of experts to provide estimates.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
40	Performance	ISO24028	Is the data used for the development, training and implementation of the algorithmic system representative for the task (this includes checks to see if there is enough data and if the populations are represented fairly)?	Controls are in place to ensure or give reasonable assurance that sufficient data has been provided/used for the algorithmic system to generate accurate results that are fair towards all populations affected or represented by the algorithmic system and the data being used.
41	Performance		Is the data the algorithmic system is using, derived from other models?	Review whether separate algorithmic systems also need to be part of the scope of the review.
42	Performance		What processes does the algorithmic system use to handle input data?	• Review how input data is included in the algorithmic system, this could include considerations such as how data is cleaned or transformed from the original source, and how easily this is repeated when the system is refreshed; • Check that data is applied consistently throughout the algorithmic system.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
43	Performance	ICO	Is the lineage of the data used for the development, training and implementation of the algorithmic system documented so that sources, changes and alterations can be traced?	A data lineage map is built, including data origin, data stops along the way, and an explanation on how and why the data has moved over time. In case data lineage maps are not available, algorithmic system output is decomposed into data elements. Data lineage documentation should include at a minimum: • Identification of the "golden"/ authoritative source; • Description of each stage of the data sourcing process from golden source to AI-solution (e.g., transformations, filtering, sampling, manual interventions); • Analysis of the impact of the data sourcing process on AI-solution input, training and/or runtime data (e.g., reliability, assumptions, limitations and/or weakness). Data lineage maps are periodically reviewed by data owners.
44	Security	ISO27001 A.9.1	Has a (data) access control policy been established, documented and reviewed based on AI security requirements?	• Check whether access control policy has been established. • Check whether the algorithmic systems and data has been addressed in this policy.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
45	Security	ISO27001 A.9.1	Has a process of user management been implemented on data and the algorithmic system?	Check existence of users and access rights registration like access control lists (ACL's). Check whether access to algorithmic systems en data has been covered by these ACL's.
Data Preparation				
46	Ethics	EC/AI HLEG April 2019 chapter II 1.3	Did the organization establish oversight mechanisms for data collection, storage, processing and use?	Document in the PIA all types of data, how they are collected, processed, stored and used in deployment. The PIA should also contain the volume, variety and sensitivity of the data.
47	Security	EC/AI HLEG April 2019 chapter II 1.3	Did the organization align the system with relevant standards (for example ISO, IEEE) or widely adopted protocols for daily data management and governance?	Investigate (by inquiry of the system owner or inspection of system documentation) whether the algorithmic system has been designed consistent with generally accepted standards for information security? Obtain and inspect the system design documentation.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
48	Privacy	EC/AI HLEG April 2019 chapter II 1.3	What protocols, processes and procedures did the organization follow to manage and ensure proper data governance?	• Did the organization assess who can access users' data, and under what circumstances? • Did the organization ensure that these persons are qualified and required to access the data, and that they have the necessary competences to understand the details of data protection policy? • Did the organization ensure an oversight mechanism to log when, where, how, by whom and for what purpose data was accessed?
49	Privacy		Has data been de-identified where possible?	Identify the use of de-identification software.
50	Performance	COBIT	Is the data correctly annotated for the intended purpose of the algorithmic system? Is the annotation process documented correctly, reperformable and unbiased?	Check whether an ontology exists for specific domain vocabulary, formatting and categorization of the data used. Data should be annotated correctly and completely based on the type of data and by domain experts, all of which is reviewed for completeness and accuracy.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
51	Security	ISO27001	• Has a process been implemented to ensure the quality and integrity of the data? • Does this process verify that data sets have not been compromised or hacked? • Has this process also addressed input validation and integrity checks on user-supplied input?	Check existence of validation checklist and outliner analysis.
52	Security	ISO27001	Has well-formed input been defined?	Check existence of validation checklist and outliner analysis.
53	Security	ISO27001 A.9.12	When training takes place against online data stores, how secure is the connection between the algorithmic system and the data?	Check network configuration parameters, and verify data communication between the algorithmic system and the data source has been secured.
54	Security	ISO27001 A.13.2	Have all used data sources been verified?	• Check whether the data source authenticity has been verified on network level (network configuration parameters); • Check whether third-party assurance has been provided on data quality in the data source.
55	Security	ISO27001 A.15	When data is sourced from a third party, has it been ensured that this third party has strong security practices?	Check whether third-party assurance has been provided on data quality in the data source.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
56	Security		Can malicious input data be detected?	Check whether data analyses have been performed on data quality.
Modeling				
57	Ethics	EC/AI HLEG April 2019 chapter II 1.1	How does the algorithmic system interact with decisions by human (end) users (e.g., recommended actions or decisions to take, presenting of options)?	Has the organization documented: • Whether the algorithmic system should communicate such decision, content, advice or outcome is the result of an algorithmic decision? • In case of a chatbot or other conversational system: ○ Whether the human end users are made aware that they are interacting with a non-human agent; ○ Whether the AI-system could affect human autonomy by interfering with the (end) user's decision-making process in an unintended way?

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
58	Ethics	EC/AI HLEG April 2019 chapter II 1.3	Is there a self-learning or autonomous algorithmic system or use case? If so, did the organization put in place more specific mechanisms of control and oversight?	Has the organization documented which detection and response mechanisms are established to assess whether something could go wrong? For example, a stop button or procedure to safely abort an operation where needed, and is this procedure to abort the process entirely, in part, or delegate control to a human?
59	Ethics	EC/AI HLEG April 2019 chapter II 1.4	Can the organization provide an explanation as to why the algorithmic system took a certain choice resulting in a certain outcome that all users can understand?	Verify whether the organization has obtained legal advice and involvement of the DPO to meet the legal and privacy requirements are met.
60	Ethics	general legal	How does the organization measure, mitigate and monitor regularly how the model performs against prohibited discrimination grounds?	For example, a confusion matrix can support in measuring the performance of the model, compared to a neutral set of data that contains no unlawful bias.
61	Privacy	GDPR, chapter 2	• Have all the purposes (including secondary purposes) of the algorithmic system been identified? • Are all direct and indirect types of personal data of the algorithmic system identified? • Is there a lawful basis for all the purposes of the algorithmic system where it concerns personal data?	Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the assessment of purpose, personal data categories, and the lawful basis for processing.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
62	Privacy	GDPR, chapter 4	Has a DPIA been carried out to assess the data protection of the algorithmic system?	Verify that a Privacy Impact Assessment (PIA) has been performed and documented for the algorithmic system with the necessary depth and substance to meet the requirements of laws and regulations, including the assessment of data protection measures.
63	Performance		Is there a profound understanding of the algorithmic system?	• Availability of a simple design/picture and/or description representing the AI algorithmic system; • Are inputs, calculations and outputs separately identified?
64	Performance	EC/AI HLEG April 2019 chapter II 1.2	To what degree could the algorithmic system be dual-use? If so, identify if suitable preventative measures are taken.	For example, not publishing the research or deploying the system.
65	Performance	EC/AI HLEG April 2019 chapter II 1.2	Has the organization ensured that the algorithmic system has a sufficient fallback plan if it encounters adversarial attacks or other unexpected situations?	For example, technical switching procedures or asking for a human operator approval before proceeding.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
66	Performance	EC/AI HLEG April 2019 chapter II 1.2	Is a strategy in place to monitor and test if the algorithmic system meets the goals, purposes and intended applications? IS tested whether specific contexts or particular conditions need to be taken into account to ensure reproducibility? Are verification methods to measure and ensure different aspects of the system's reliability and reproducibility in place? • Are processes to describe when an algorithmic system fails in certain types of settings in place? • Are these processes clearly documented and operationalized for the testing and verification of the reliability of algorithmic systems? • Are mechanisms of communication established to ensure (end)users of the algorithmic system's reliability?	*Intentionally left blank*
67	Performance	EC/AI HLEG April 2019 chapter II 1.6	Are mechanisms established to measure the environmental impact of the algorithmic system's development, deployment and use	For example, the type of energy used by the data centers.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
68	Performance		Does the model respond logically to basic changes being made to the algorithmic system inputs?	Review how changing basic algorithmic system inputs impact the outputs, for example by: • Simplifying settings to the most basic scenario; • Investigating the initial (starting) conditions for the model; • Sensitivity analysis with realistic input variations; and • Sensitivity analysis with extreme or implausible inputs variations.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
69	Performance		How accurate is the detail of the algorithmic system?	• Take sample checks to assess whether the model is doing what it should, for example by re-performing calculations on sections of the model; • Consistency of accuracy and aggregation of the data; • For Excel based models identify areas that might expose weaknesses in the model, such as: ○ Circular reference warnings; ○ Hard coding of values; ○ Linking of data from other files; ○ Complexity of formulae. • For syntax-based models, review whether comments or notes explain what the element of the model is doing and whether it is understandable to someone unfamiliar with the model.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
70	Performance		Are the details of algorithmic system assumptions recorded and justified?	• Identify and review the list of assumptions, for example: ○ Suitability of selection based on the purpose of the model; ○ Underlying evidence – source and quality; ○ Level of simplification/complexity; ○ Rationale for level of accuracy and aggregation; ○ Distinction between data and structural assumptions; ○ What are the main assumptions in the model?
71	Security	ISO27001 A.4.14	• Has training data and systems that host them been part of an AI risk and threats assessment (A.4.14, A.14.3)? • Did you assess potential forms of attacks to which the AI-system could be vulnerable? • Did you consider different types and natures of vulnerabilities, such as data pollution, physical infrastructure, cyberattacks?	Identify risk assessments performed and verify that the algorithmic systems and training data are in scope.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
72	Security	ICO	Are precautions and checks in place to safeguard the interaction between the algorithmic system and other entities that could alter or corrupt input or output data?	Identify and monitor data transfer between algorithmic system(s) or entities to detect indications of compromised appropriateness (ideally through automation). Where such a compromise is detected, take appropriate action. Where the AI-systems are in an IaaS or a PaaS environment ensure that the service provider has appropriate controls in operation and that compromises are reported promptly and fully.
73	Security	ISO27001 A.14.2.6	• Does the algorithmic system development and training environment meet the AI-security requirements (A.14.2.6)? • Did measures or systems being implemented to ensure the integrity and resilience of the algorithmic system against potential attacks? And is being verified how you're the system behaves in unexpected situations and environments?	• Check whether AI-security requirements are documented for AI-system development and training environments; • Check whether measures are taken to protect de AI-system against cyberattacks. Check whether tests has been performed on unexpected situations; • Check whether system monitor measures are implemented.
74	Security	ISO27001 A.9	Does the algorithmic system development and training environment being protected by an access control system?	Identify existence of access control system and access control lists (ACL); Identify existence of an access control process.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
75	Security	ISO27001 A.15	When using pre-built algorithmic systems from third parties, has the quality of this models and its providers being verified?	Check whether third-party assurance has been provided on third-party algorithmic systems.
76	Security	EC/AI HLEG April 2019 chapter II 1.7	Did you establish processes for third-parties (e.g., suppliers, consumers, distributors/vendors) or workers to report potential vulnerabilities, risks or biases in the algorithmic system?	Check whether a process is being implemented in order to detect system and software vulnerabilities, security risks or bias.
Evaluation				
77	Ethics	EC/AI HLEG April 2019 chapter II 1.2	Did the organization estimate the likely impact of a failure of your algorithmic system when it provides wrong results, becomes unavailable, or provides societally unacceptable results (for example discrimination)?	Investigate whether the organization has: • Established an adequate set of mechanisms that allows for redress in case of the occurrence of any wrong, harm or adverse impact; • Defined thresholds and put governance procedures in place to trigger alternative/fallback plans; • Defined and test fallback plans.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
78	Ethics		On which aspects does the organization monitor the algorithmic systems?	Verify if the monitoring of algorithmic systems take place on multiple levels, namely: • Data monitoring (data science issues); • Prediction monitoring (data science issues); • System monitoring (operational issues).
79	Performance		How accurate does the algorithmic system perform against historical data?	• Review (or perform) checks assessing how the algorithmic system predicts known history, both on data available during development and since implementation; • For older algorithmic systems, use back casting to determine its 'forecasting' record.
80	Performance		Has the algorithmic system been subject to external review during or after development?	• Identify who has reviewed the algorithmic system, and why; • Review documentation produced by bodies reviewing the algorithmic system. This is not limited to the building of the model and could cover any of the areas outlined in this framework; • Identify whether there is an external assurance statement from experts.
81	Performance		Has the status of the assumptions been critically compared to third party sources, or benchmarked against industry norms?	• Check to similar algorithmic systems; • Check to published standard assumptions.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
82	Performance		What are the uncertainties of the algorithmic system?	• Review whether uncertainty has been quantified in the model (i.e., are high and low estimates provided alongside a point estimate?); • Review whether the model estimates the level of confidence in the output; • In the context of materiality, consider developing: ○ a list of modelling uncertainties; ○ a list of input data, evidence and intelligence used in the analysis and consider each type of uncertainty that could affect it; • Diagram representing key parts of the model with consideration for what additional factors might act at that point and affect the analysis outcome.
83	Performance		Has a sensitivity analysis been performed to calculate ranges or the likelihood of outcomes occurring?	• Review whether levels used in sensitivity analysis are realistic and conservative based on the source data; • Review or perform analysis such as Monte Carlo simulation or scenario analysis.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
84	Performance		Are the assumptions/hypotheses/correlations formulated between the input data and the output from the algorithmic system correct and checked to prevent mistakes in correlation or causality?	Quality controls exist to help ensure the appropriate relationships between variables/ events and hypotheses are defined. This includes interdependencies and distinguishing correlation and causality (e.g., through Bayesian statistics, Hybrid Monte Carlo methods, or causal models such as Granger non-linear causality, Neyman-Rubin, Pearl and/or Granger).
85	Performance		Do changes in the inputs/assumptions have a material or significant impact on outputs?	• Review or perform additional runs of the algorithmic system to test sensitivities on outputs when the assumptions are changed; • Review or perform additional runs of the algorithmic system to test sensitivities on outputs when inputs are changed.
86	Performance		Have issues over poor-quality data and assumptions and other identified risks been addressed?	Test for the impact of weak information in the algorithmic system.
87	Performance		Are you able to validate the algorithmic system outputs?	Review appropriateness of algorithmic system output by comparing to: • Previous runs of the model; • Other models such as parallel systems; • Independent sources.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
88	Performance		Are decisions based on the algorithmic system output proportionate to the robustness of the model?	Review whether decisions are appropriate and proportionate to the robustness of the algorithmic system, for example considering monetary impact of decision given constraints of the algorithmic system.
89	Performance		Is the output from the model adjusted outside of the algorithmic system?	Review whether any additional procedures or adjustments that are made to the algorithmic system output are justified and how they impact on the robustness of decisions made.
90	Performance		Does the model output meet the requirements and aims of the algorithmic system as outlined in the algorithmic system concept?	Compare the actual outputs of the algorithmic system with the aims of the concept.
91	Performance	ISO24028	Are dynamic learning algorithmic systems being correctly monitored to ensure operation within acceptable limits and so as to prevent undesirable and runaway behavior?	The risk of runaway outcomes is assessed. In case there is a risk of runaway outcomes, either mitigating controls are implemented to suspend AI activity or the absence of mitigating controls is justified. Examples of mitigating controls are: kill switch, fallback, revert to previous iteration, limits on volume or types activities that can be performed by the algorithmic system.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
92	Performance		Have (KPIs and KRIs) metrics been defined to monitor the algorithmic system's performance and are these adequate?	To determine the algorithmic system's performance statistical metrics commonly used. Examples are: • Classification accuracy; • Logarithmic loss; • Area Under the Curve (AUC); • Precision-Recall Curve; • Receiver Operating Characteristics (ROC). These metrics can be used to calculate a Confusion matrix.
93	Performance		Have different methods/approaches been selected to evaluate the algorithmic system performance?	There should be a proper mix of methods/approaches for evaluating and testing. Examples which be used are: • Formal statistical metrics; • Empirical testing (e.g., benchmarking, expert panel); • Human verification: are the algorithmic system results not worse than the human intelligence results? • Field trials, e.g., chatbots.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
94	Performance		Is there a way to track the deployed algorithmic system's version?	Check whether model versions are tracked. Configuration errors might occur if different versions of the algorithmic system are used.
95	Performance		Is there a periodic review with stakeholders to identify any significant missing items and is reasonableness of targets and tolerances redefined?	If no periodic review with key stakeholders takes place, identify the rationale for why not.
96	Performance		Is there a clear dashboard available which shows performance results that are easy to understand for stakeholders?	Points to focus on are: • How does the AI-system function? Are the results made understandable by the processing of input features (causality); • Is the societal context made clear? What are the relevant regulations, standards and organizational processes with regards to the AI implementation; • Is the required explanatory power of the algorithmic system identified? This can differ based on e.g., the impact of AI decisions, the usage of personal sensitive data and the needs of the specific stakeholders.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
97	Performance		Are the results from the algorithmic system presented correctly and understandably so as to ensure all involved parties are adequately informed and are able to understand the core aspects of the algorithmic system?	Failure to provide adequate information to algorithmic system stakeholders, to enable other individuals to understand the core aspects of the algorithmic system design (e.g., trace outputs to inputs), leading to: • (For developers and validators) inability to challenge or replicate the algorithmic system; and/or • (For users) misuse of algorithmic system outputs.
98	Performance		Based on which interval is the algorithmic system's performance revaluated?	Request if there is a formalized procedure regarding the periodic revaluation for models.
99	Performance		• Is an override process in place for exceptions (controllability)? • Is a root cause analysis performed for exceptions that are deemed incidents?	An example of an override is an emergency stop "button" or automated stop and hold until released after human interaction.
100	Performance		Is a process in place to assess exceptions in the algorithmic systems performance?	Are exceptions assessed for risk and performance against risk appetite and business impact/ criticality, and take inter-dependencies in to consideration. If no assessment takes place, identify the rationale for why not.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
101	Security	ISO27001 A.1.2.4	Has a process been implemented in order to detect input attacks and poisoning of training data?	• Check whether access to (training) data has been registered in logging files; • Check whether these logging is being monitored and analyze in order to detect attacks or poisoning.
Deployment				
102	Ethics	EC/AI HLEG April 2019 chapter II 1.4	• How does the organization you communicate to (end)users – through a disclaimer or any other means – that they are interacting with an AI-system and not with another human? • Has the organization labeled the AI-system as such?	Verify the communication of the system end-users. Additional consideration should be given to communication in case the AI-system interacts directly with humans, the AI-system to encourage humans to develop attachment and empathy towards the system. If applicable, verify that the AI-system clearly signals that its social interaction is simulated and that it DPBU has no capacities of “understanding” and “feeling”.
103	Performance		Has the algorithmic system been published?	If the algorithmic system has not been published, identify the rationale for why not.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
104	Performance		What documentation and processes are in place to ensure a corporate memory for the algorithmic system exists?	- Review how changes to the algorithmic system, for example, detail of change, rationale and impact, are recorded; - Review the adequacy of any model documentation (technical and non-technical) provided for new users, for details of what the algorithmic system does and how to operate it.
105	Performance		What process is used to change/update assumptions?	- Review the process for managing how assumptions are changed within the algorithmic system; - Review whether assumptions should have been updated in light of any changes to circumstances.
106	Performance		What is the process for the routine review of outputs?	Review process for circulating outputs internally and externally, checks could involve different roles, for example: - Technical staff not directly involved with the model; - Senior staff responsible for the model; - External expertise.
107	Performance		Are the limitations and uncertainty of the model output communicated to decision makers?	Review how model outputs are presented to decision makers, for example how findings are presented in a business case.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
108	Performance		Are the outputs from the model responsive to the ongoing needs of the organization?	Review whether the model is being used to track on-going performance as a monitoring tool.
109	Performance		Are forecasts compared with actual outputs in order to validate the results and inform future development?	Compare the actual outputs with reality to check accuracy and check whether this is used to update future iterations.
110	Performance	EC/AI HLEG April 2019 chapter II 1.2	Did the organization put in place ways to measure whether its system is making an unacceptable number of inaccurate predictions?	Assess if a well-formed development and evaluation process is present. In case these inaccurate predictions cannot be avoided can the system indicate how likely these errors are.
111	Security	ISO 27001 A9.	Have measures been taken to ensure that only authorized users/customers/partners have access to the algorithmic system, data and output?	• Ensure privacy and confidentiality of people's data; • Ensure that models and output cannot be modified by unauthorized staff; • Ensure security risks of all externally maintained software is included.
112	Security	ISO 27001 A12.1.4	Has a separation of development, training and operational environments been implemented?	Check whether environments are being separated and unauthorized access or changes in (training) data and models are being prevented.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
Governance				
113	Roles & responsibilities		Who is the single Senior Responsible Owner (SRO) within the organization for the algorithmic system, and has the SRO approved the algorithmic system before deployment and implementation?	• Check whether documentation of roles and responsibilities throughout the algorithmic system's development, maintain and run processes is available. • Check whether the SRO, or maybe another business owner, has approved the algorithmic system before deployment and implementation in the production environment.

CRISP-DM phase/Governance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
114	Roles & responsibilities		Does the organization have sufficient documentation in place on governance and quality assurance for their algorithmic system?	• Check whether roles and responsibilities (i.e., commissioner, lead analyst, lead analytical assurer) are documented? • What processes are in place for succession planning/handover, i.e., when a key person leaves the modelling project? • Check whether the algorithmic system has been developed in collaboration with stakeholders (i.e., customers)? For example: • Are requirements captured and documented into a specification? • Are assumptions listed and agreed? • Check whether there is an agreed quality assurance plan throughout the model development process. • Check whether there is evidence a customer of an algorithmic system has influenced it to meet expectations ('gaming')?

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
115	Ethics	General legal	How does the organization regularly monitor the outcome of the algorithmic system against unlawful bias?	• Check whether there are thresholds set for acceptable outcomes of the model to avoid unlawful bias. • Check whether the model is regularly measured against the thresholds. • Check whether the outcome is reported and available for audit. • Check whether the model can be changed in case the outcomes exceed the set thresholds.
116	Privacy	GDPR, chapter 2	• Who determines the purpose and means of the algorithmic system? • If multiple, are the scopes of responsibilities defined?	Investigate if an AI System "Charter", Terms of Reference or other formal documentation exists that defines the roles and responsibilities on system and data governance.
117	Privacy	GDPR, chapter 4	Has the Data Controller implemented a risk management process on the privacy and data protection risks with documented information to demonstrate the operating effectiveness of the process?	Investigate (by inquiry of the system owner or inspection of system documentation) whether the organization applies a systematic and documented information security risk management process to the algorithmic system.

CRISP-DM phase/Gover-nance phase	Risk category	Reference to applicable laws, regulations, frameworks and standards	Key questions to consider	Examples of checks to perform and/or evidence to look for
118	Performance		Is the algorithmic system business within the organization critical?	• Define what makes an algorithmic system 'business critical'. Test this definition with definitions from other organizations; • Evidence the Accounting Officer's governance statement (typically within the annual report) includes an appropriate quality assurance framework for business-critical models; • Evidence the Accounting Officer maintains an up-to-date list of business-critical models and that this is publicly available.
119	Performance		How are algorithmic system outputs challenged and used within the organization?	• Is there a forum available for people outside the algorithmic system development process to challenge the development and use of model outputs? • How do algorithmic system customers develop an understanding of the caveats of the model? • Are algorithmic system limitations and caveats reported alongside the main outputs of the model?