The Aletheia AF Framework™ 2.0

ADAPTED FOR USE IN ONCOLOGY AI PROJECTS

AREAS	CONTEXT	ETHIC	#	REALISATION PRINCIPLES FOR ONCOLOGY	EVIDENCE
Social Impact	Benefits	AI and robotics shall be seen as delivering good. Doing good is one of the five key ethical principles of the EU guidelines for ethical AI. Good includes commercial prosperity.	1	Improve employee well-being via safety, work conditions and job satisfaction. Applies to nurses, doctors and all health professionals.	
			2	Deployment of AI shall be demonstrated to provide improved patient, professional and cost outcomes, not just "AI for the sake of AI". Where possible, formal ethnography studies should be undertaken to demonstrate this.	
	Human impact	AI systems should be used to enhance positive social change and enhance sustainability.	3	It shall be clear at what step in the disease evalutaion/mangement process AI is being deployed.	
			4	Early analysis shall be undertaken to identify the effect on patient-professional communication, as well as any potential job role changes.	
			5	Potential upskilling opportunites such as interpretation and communcations for decision support, or improved anatomy training for contouring, or accrediation in data science for radiotherapy shall be explored as soon as any effect on empoyees is established.	
			6	Analysis shall be undertaken to assess the impact of the AI deployment within the department, allied health, referring physicians, the patient and in community based care - particularly assessing the likelihood for the technology to have a negative impact on the department, the same assessment should be afforded to patients and patient representatives where appropriate.	
			7	Where there is potential for negative impact on the department, allied health, referring physicians, the patient or community-based care it shall be discussed with them when possible to allow them to adapt to any changes. This same opportunity should be afforded to patients and patient representatives where appropriate.	
	Communication	Knowledge of the human interactions with AI should be provided by key stakeholders.	8	Frequent communication and discussion should be had with all key parties – in particular medical professionals and paitient representatives – through a variety of channels. (Need to enhance without excessive alert or signal fatigue).	
	Loss of skills	AI systems should be used to enhance positive social change and enhance sustainability.	9	The AI system must improve the therapeutic relationship and team collaboration. Analysis also shall be undertaken as to whether any loss/reduction of skills needs to be sustained, as well as the cost effectiveness.	
Accuracy/Trust	Safety	AI systems should be safe and secure throughout their operational lifetime. This should be verified where applicable and feasible.	10	A risk analysis shall be undertaken with specific emphasis given to identifying and mitigating any hazards to human safety.	
	Transparency and traceability	AI systems must provide for transparency and traceability of their design, inputs and outputs.	11	The algorithms shall be assessed for any bias or discrimination impact and their provenance shall be clearly stated to enable any future Root Cause Analysis or troubleshooting (Note, for complex systems, it may be difficult to assess the risk of bias. A new bias assessment template has been created as part of an ecosystem of AI Assurance tools at [link]).	
			12	The provenance of all training data shall be clearly stated to enable any future RCA.	
			13	The hierarchy of decision making shall be clearly stated regarding human v AI.	
			14	It shall be clear what the insight (forecast/decision making etc.) improvement is compared with a human – forecast improvement and actual.	
	Bias	AI systems must be free from bias or prejudice.	15	It shall be clearly stated how any training data sets or nomograms have been assured to have no unintentional or unethical biases. Need to be transparent about the generalizability of nomograms and how this may effect patient outcomes.	
	Validity and reliability	For AI to succeed it must be trusted.	16	A sense check of the reslts - called an Executive Monitor - shall be depolyed in the system to determine if the outputs from the AI are as expected.	
			17	A continuous automated monitor shall be deployed in the system to continuously test the system by using existing test/synthesised data, which already has known and approved outputs.	
			18	An Independent check shall be deployed in the system – assessment of the same data using a completely independent assessment mechanism which is already approved – **Validation/explainability Check**.	
			19	A Process comprehensiveness check shall be deployed in the system – have the right number of assessments taken place?	
			20	A Faultless transmission of data shall be deployed in the system – use of Cyclic Redundancy Checks for systems where data is transmitted in a novel manner. Otherwise, use Digital Imaging and Communications in Medicine (DICOM) protocols and secure transmission protocols where appropriate.	
	Sparse data interpolation	For AI to succeed it must be trusted.	21		
Governance	Data protection	For AI to succeed it must be trusted.	22	It shall be stated whether there is, or will be, any Personal data or not.	
			23	The legitimate purpose for using the personal data shall be declared and confirmation provided that this has been agreed with the patient/legal guardian.	
			24	The architecture of the system shall protect the data from unwanted access without permission – complying with the principle of 'privacy by design and by default'.	
			25	The architecture of the system shall have the facility to, on demand, identify an individual's personal data and update, amend or remove every trace in line with privacy requirements and individuals' rights.	
			26	No personal data shall be sent outside of the relevant, legal zone (e.g. European Economic Area, US).	
	Export control	For AI to succeed it must be trusted.	27	N/A	
	Confidential information	For AI to succeed it must be trusted.	28	All confidential information shall be declared to, discussed with and the architectural protections by an IT security expert.	
	Cyber security	For AI to succeed it must be trusted.	29	All systems shall be assessed and approved by an IT Security expert.	
	Accountability	Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes.	30	Ultimate accountability for the outcomes of the AI system needs to be clearly stated.	
	Responsibility for decisions	Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes.	31	Algorithmic accountability should fall jointly on the developer and tester, or the DevOps team. They shall clearly state how they have assured confidence in the performance of their individual aspects of the AI system.	
	Risks from re-use/transfer across processes	For AI to succeed it must be trusted.	32	Any transfer of technology shall be fully risk assessed by risk analysis of technology transfer (in addition to that in Realisation Principle 10) and identifying the major severity effects and causes, along with the detectably mechanisms for the proposed controls – which shall be formally reviewed before proceeding.	