

Semantic Models in Integrating AI Regulations and Standards

Keynote: International Workshop on Artificial Intelligence Technologies for Legal Documents (AI4LEGAL) at ISWC'22

Dave Lewis

dave.lewis@adaptcentre.ie

*Deputy Director: ADAPT SFI Centre for Digital Content Technology,
Trinity College Dublin - <https://www.adaptcentre.ie/>*

Coordinator: PROTECT MSCA ITN - <https://protect-network.eu/>



A World
Leading SFI
Research
Centre



HOST INSTITUTION



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

HOST INSTITUTION



PARTNER INSTITUTIONS



University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath
Ireland's Global University



MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University



TUS

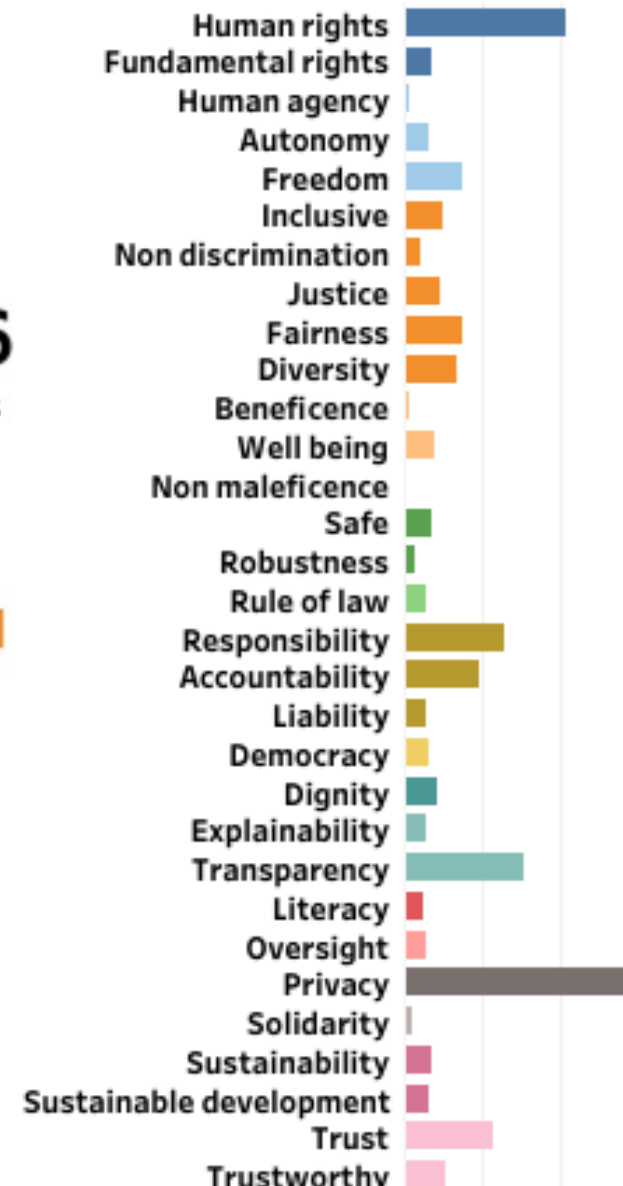
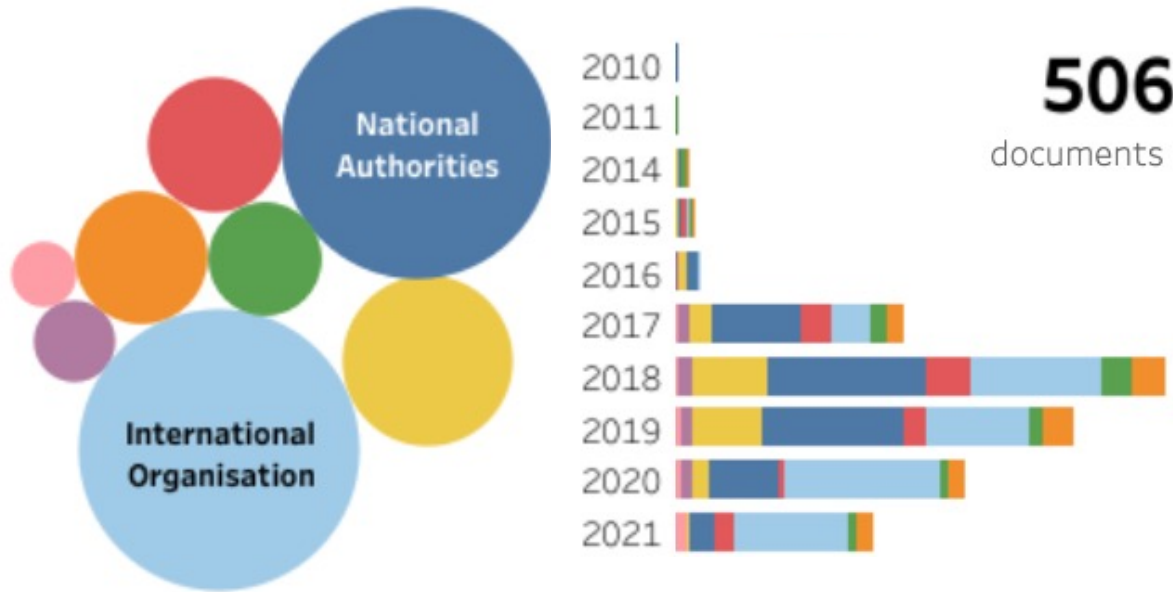
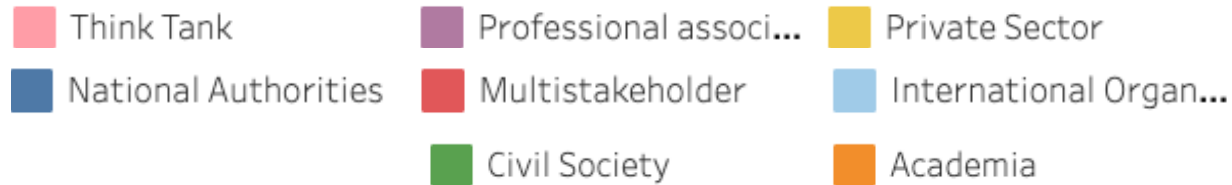


Maynooth University
National University of Ireland Maynooth



O'É Gaillimh
NUI Galway

Growth in Societal concern about Trustworthy/Responsible/Ethical AI



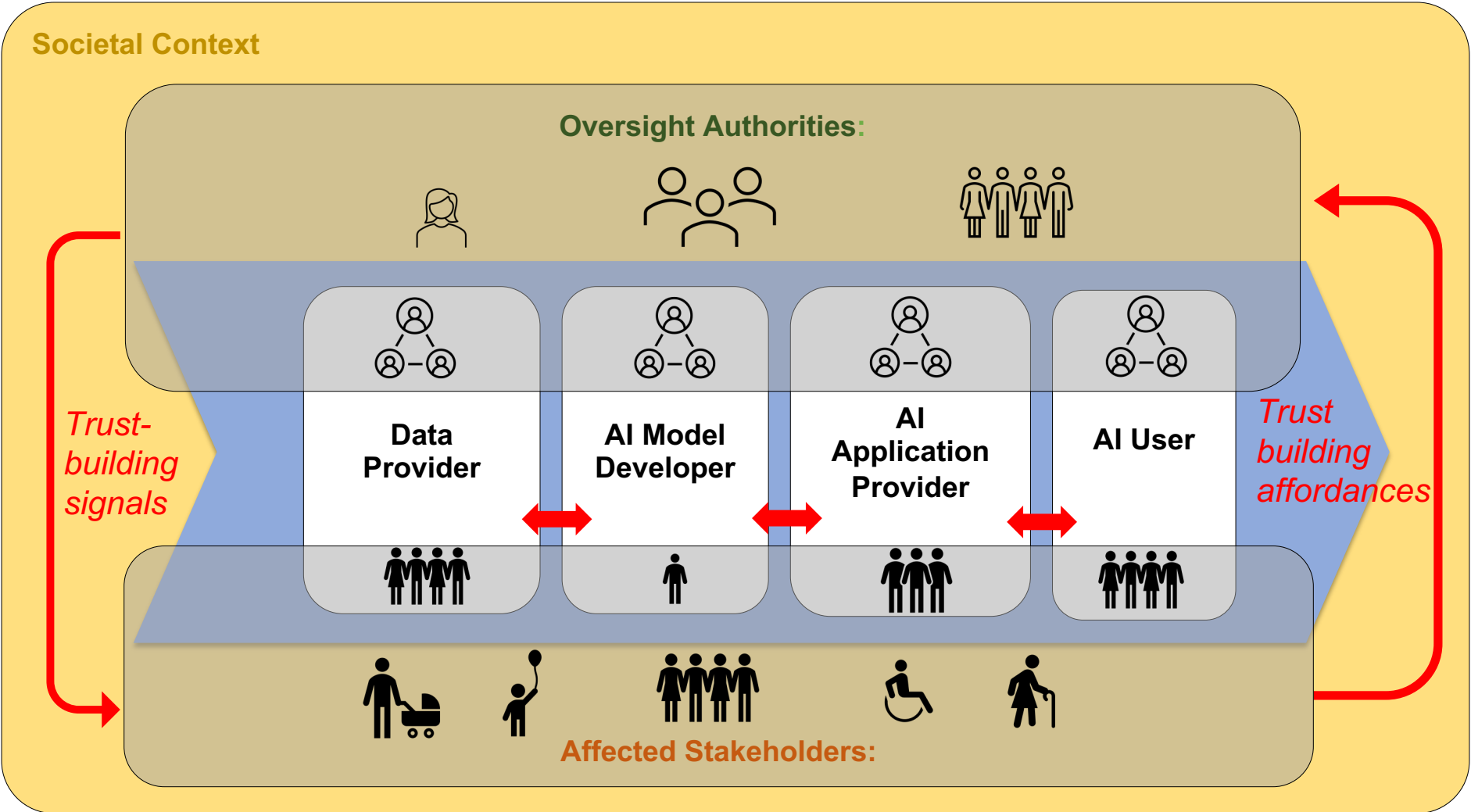
<https://www.coe.int/en/web/artificial-intelligence/national-initiatives>

Trustworthy AI and Data Governance:

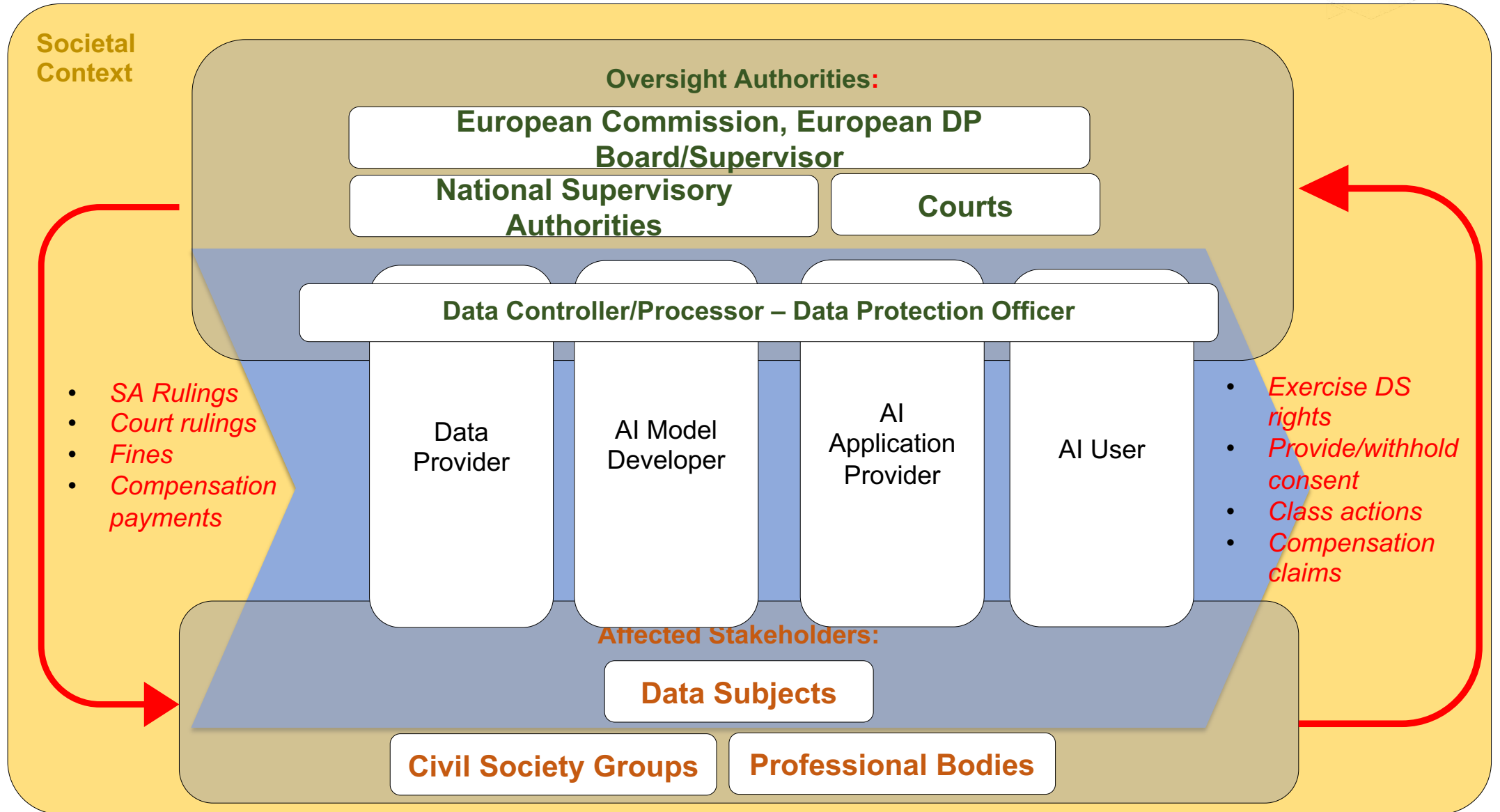
complexities moving from principles to practice:

- Significant initiatives:
 - Emerging Legislation: EU AI Act, US National AI Initiative Act
 - Emerging Standards: ISO/IEC JTC1 SC42, CEN/CLC JTC 21, IEEE P7000, National Bodies
 - Public Policy: EU HLEG, OECD, UNESCO
 - Organization level policies: tech platforms, public sector, sectoral/vertical, researchers
- Systemic Concerns
 - Workable guidelines
 - Regulatory load vs. benefits
 - Value-chain complexities & liability
 - Who wields oversight authority
 - Stakeholder engagement
 - Ethics Washing
- Divergent pressures
 - Disciplinary: 'problemists' vs 'solutionists'
 - Sectorial
 - Technological
 - Jurisdictional

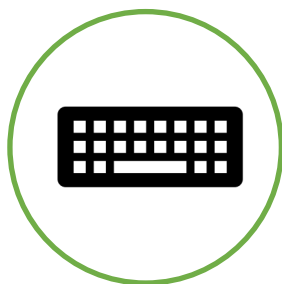
Modelling Context and Information Flows for Trustworthy AI and Data Governance



Example: Trustworthy AI/Data governance based on GDPR



Research in Semantic Modelling of EU GDPR legislation



Goals

- Open data for GDPR compliance
- Open provenance, queries, constraints



Output

- Semantic GDPR models
- Legal Text KG
- Consent
- Data Processing Provenance



Impact

- W3C Data Privacy Vocab
- Consent Receipt ISO/IEC TS 27560

Data Privacy Vocabulary (DPV)

version 0.9

[Draft Community Group Report 22 October 2022](#)

Latest published version:

<https://w3id.org/dpv>

Latest editor's draft:

<https://w3c.github.io/dpv/>

Editor:

[Harshvardhan J. Pandit](#) (ADAPT Centre, Trinity College Dublin)

Former editor:

[Axel Polleres](#) (Vienna University of Economics and Business)

Authors:

Axel Polleres (Vienna University of Economics and Business)

Beatriz Esteves (Universidad Politécnica de Madrid)

Bert Bos (W3C/ERCIM)

Bud Bruegger (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)

Elmar Kiesling (Vienna University of Technology)

Eva Schlehahn (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)

David Hickey (Dublin City University)

Fajar J. Ekaputra (Vienna University of Technology)

Georg P. Krog (Signatu AS)

Harshvardhan J. Pandit (ADAPT Centre, Trinity College Dublin)

Javier D. Fernández (Vienna University of Economics and Business)

Julian Flake (University of Koblenz-Landau)

Mark Lizar (OpenConsent/Kantara Initiative)

Paul Ryan (Unipharm PLC)

Piero Bonatti (Università di Napoli Federico II)

Ramisa Gachpaz Hamed (Trinity College Dublin)

Rigo Wenning (W3C/ERCIM)

Rob Brennan (University College Dublin)

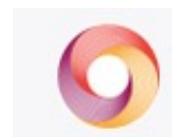
GDPRtEXT - GDPR as a Linked Data Resource

Harshvardhan J. Pandit¹, Kaniz Fatema², Declan O'Sullivan¹, and Dave Lewis¹

¹ ADAPT Centre, Trinity College Dublin, Dublin, Ireland,
firstname.lastname@adaptcentre.ie

² University of Derby, Derby, United Kingdom,
k.fatema@derby.ac.uk

Abstract. The General Data Protection Regulation (GDPR) is the new European data protection law whose compliance affects organisations in several aspects related to the use of consent and personal data. With emerging research and innovation in data management solutions claiming assistance with various provisions of the GDPR, the task of comparing the degree and scope of such solutions is a challenge without a way to consolidate them. With GDPR as a linked data resource, it is possible to link together information and approaches addressing specific articles and thereby compare them. Organisations can take advantage of this by

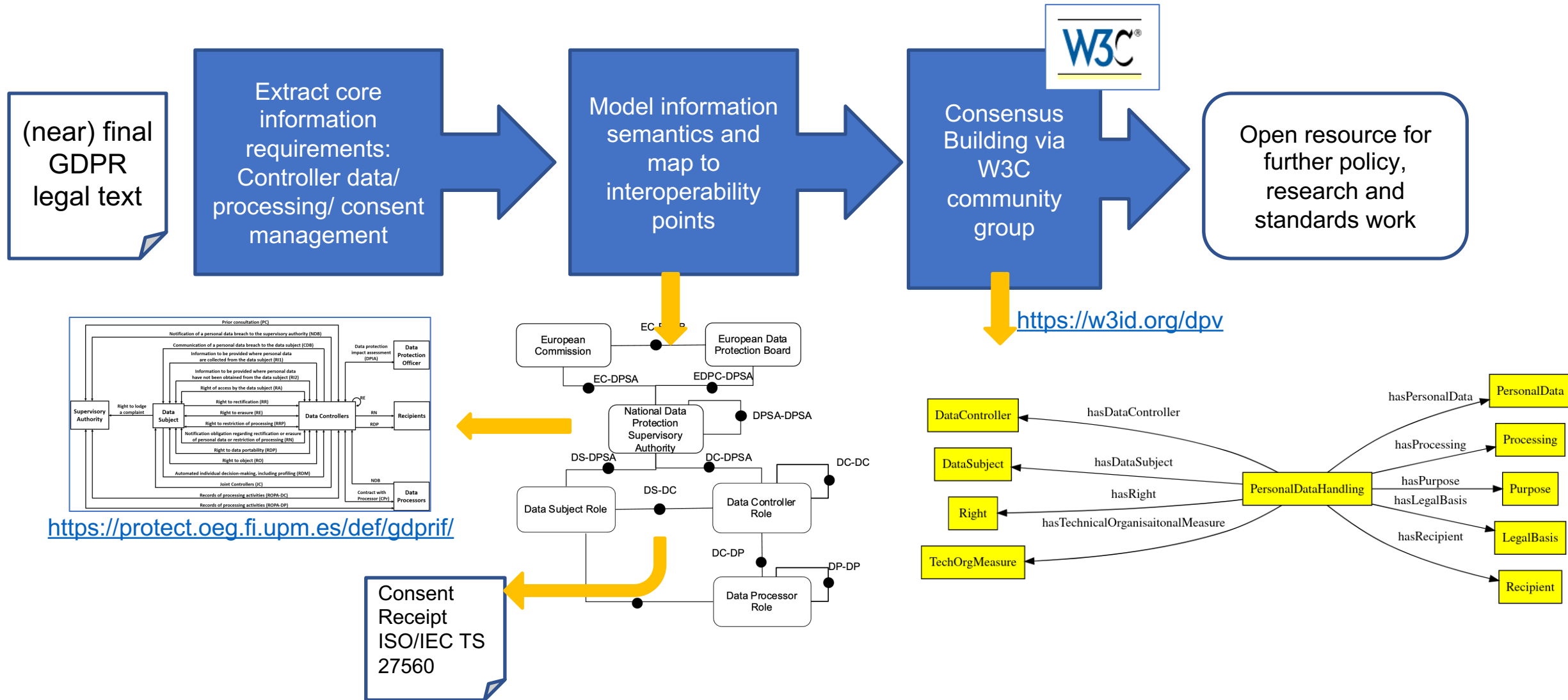


Protect ITN

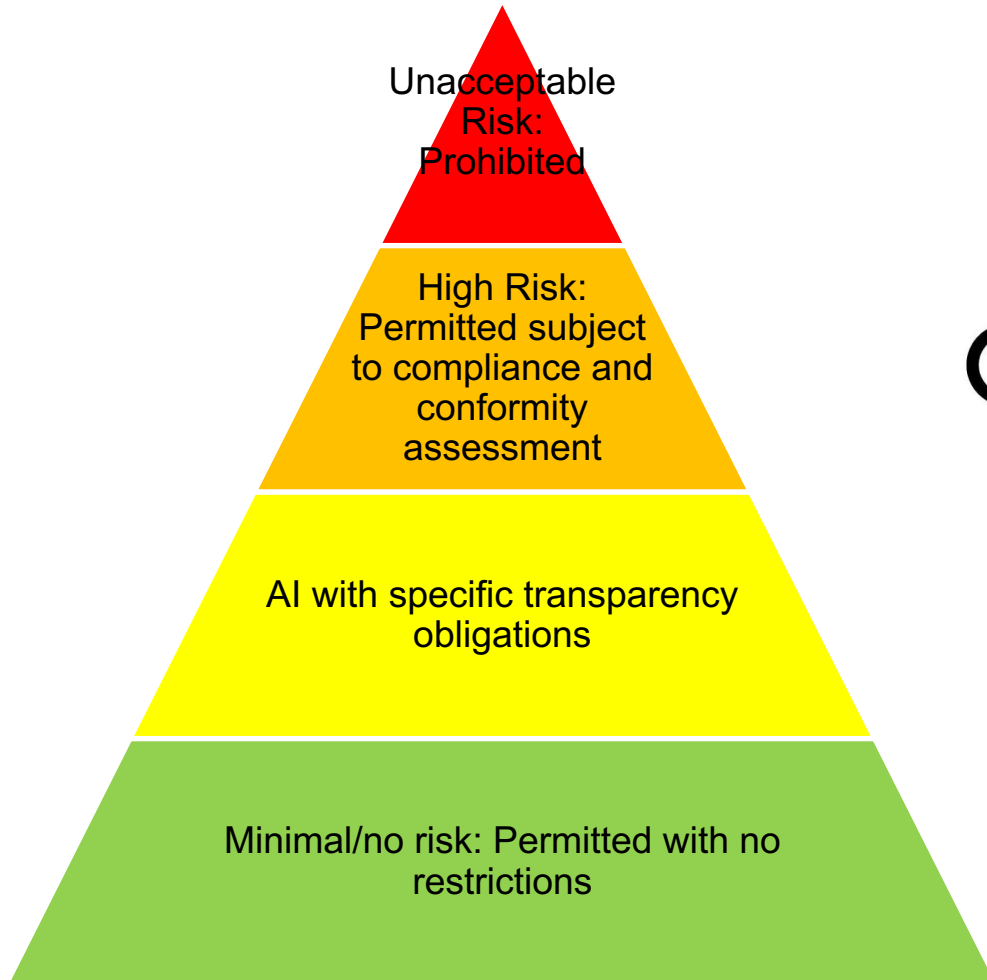


IRISH RESEARCH COUNCIL
An Chomhairle um Thaighde in Éirinn

Legal-to-Standards Mapping: using semantic models for GDPR

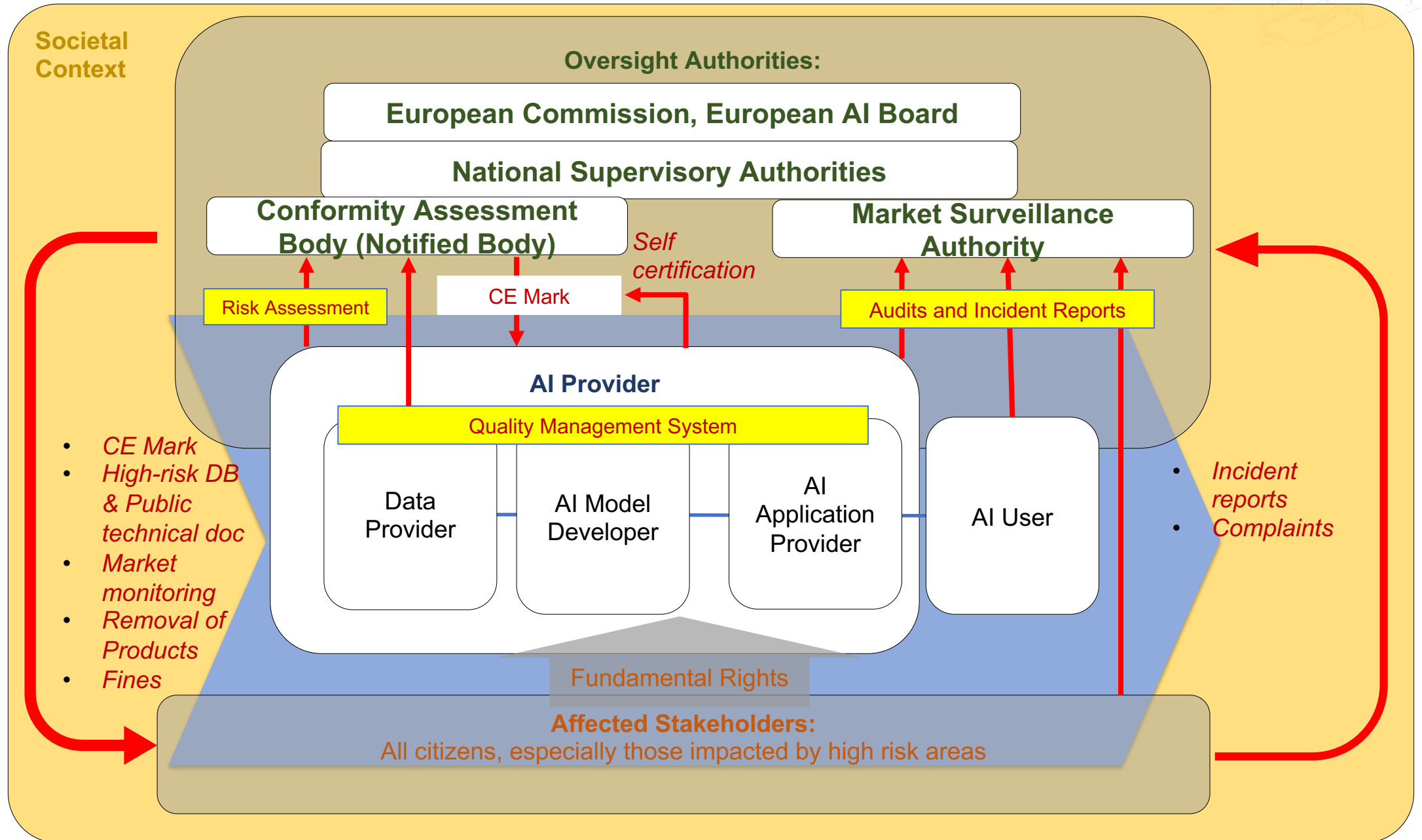


Expanding Regulatory Environment: EU AI Act

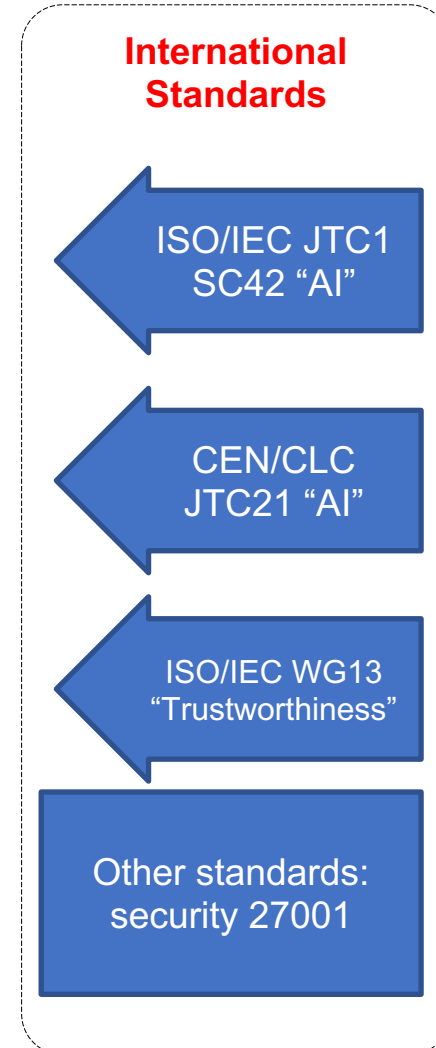
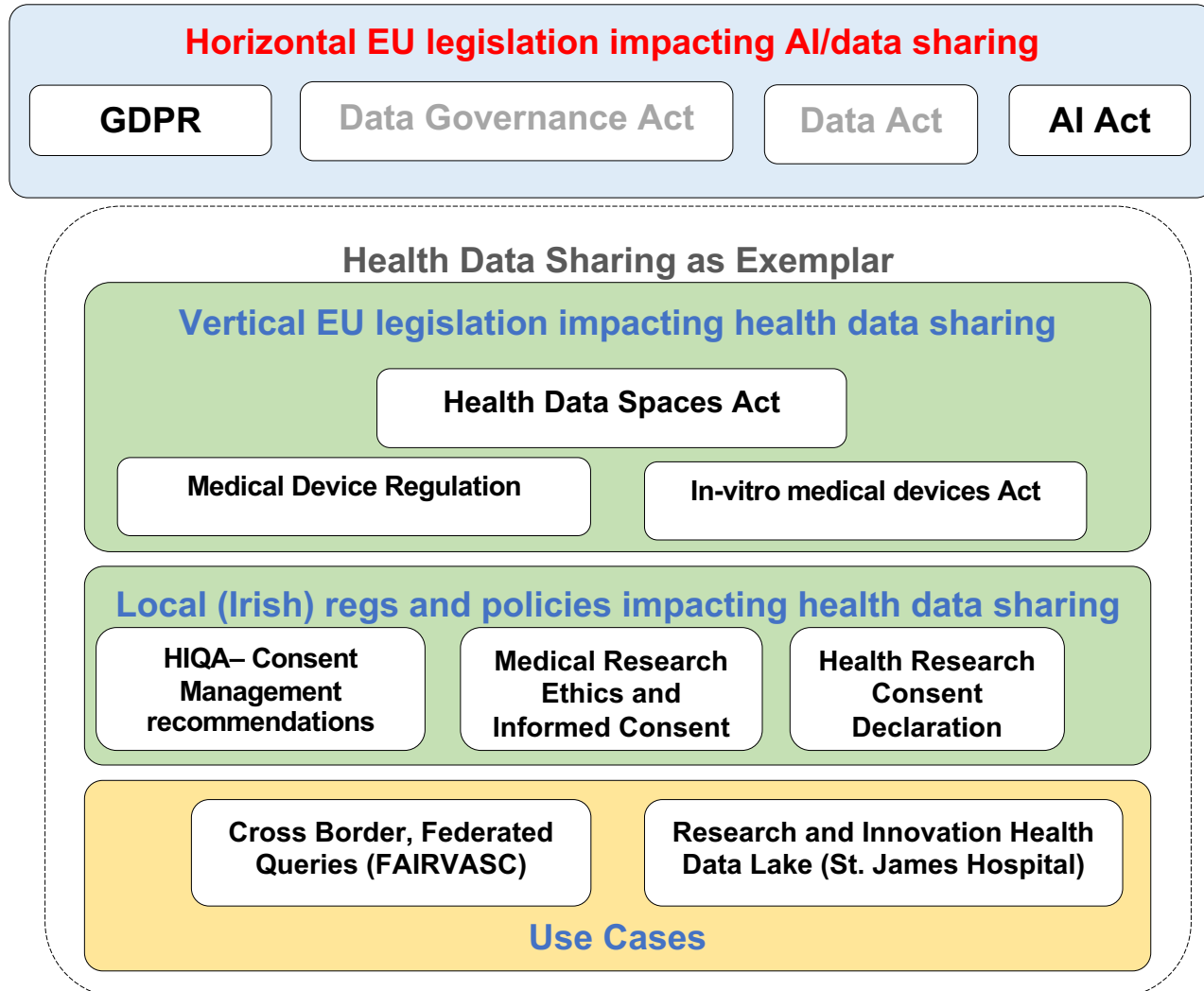


- Part of New Legislative Framework for product health and safety harmonization across single market
- Aims to protect **health, safety and fundamental rights**
- Enable access to **EU single market** for AI products/services
- A **Risk-based** approach to regulating AI
- Require product certification for high-risk AI system

Example: Trustworthy AI/Data governance based on AI Act

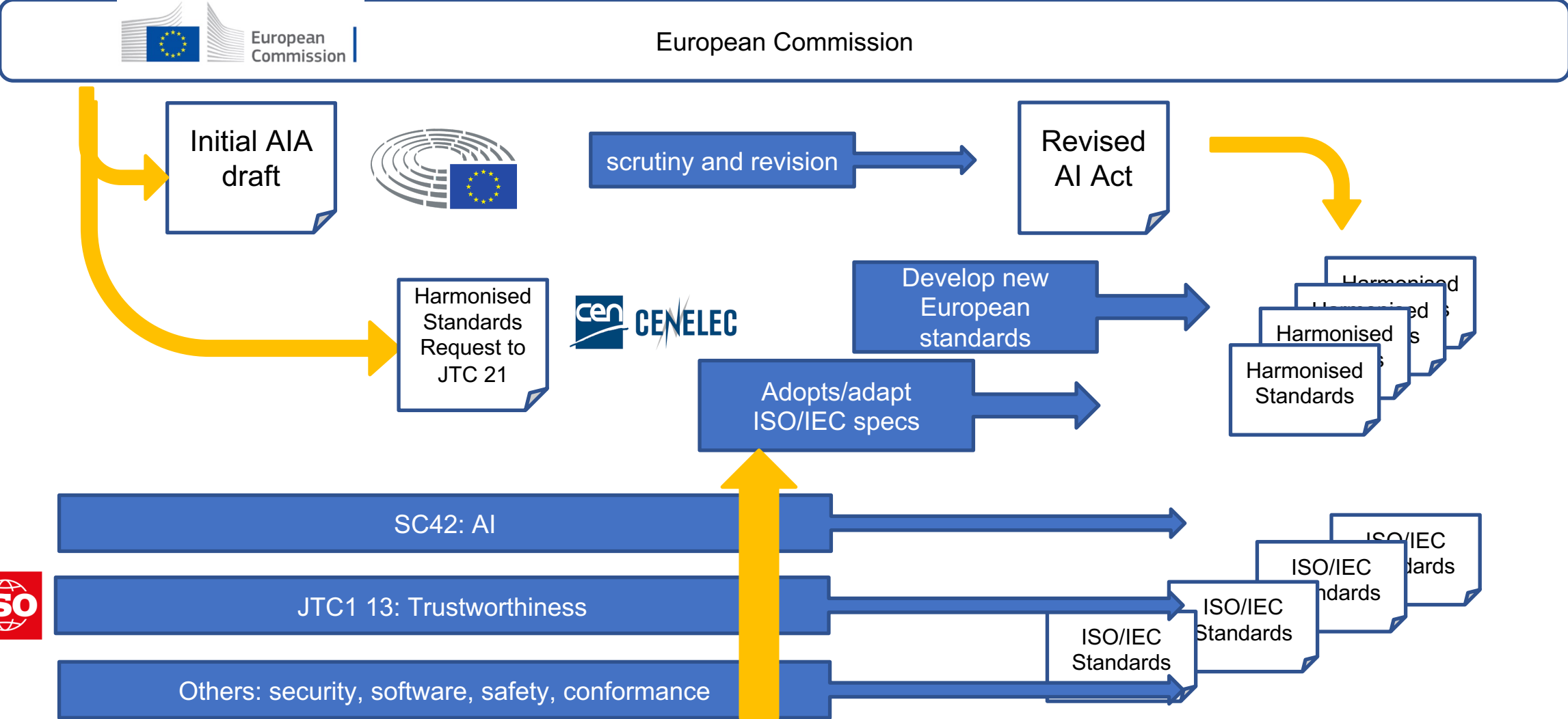


Need for Regulatory to Standards Mapping: e.g. for health data sharing

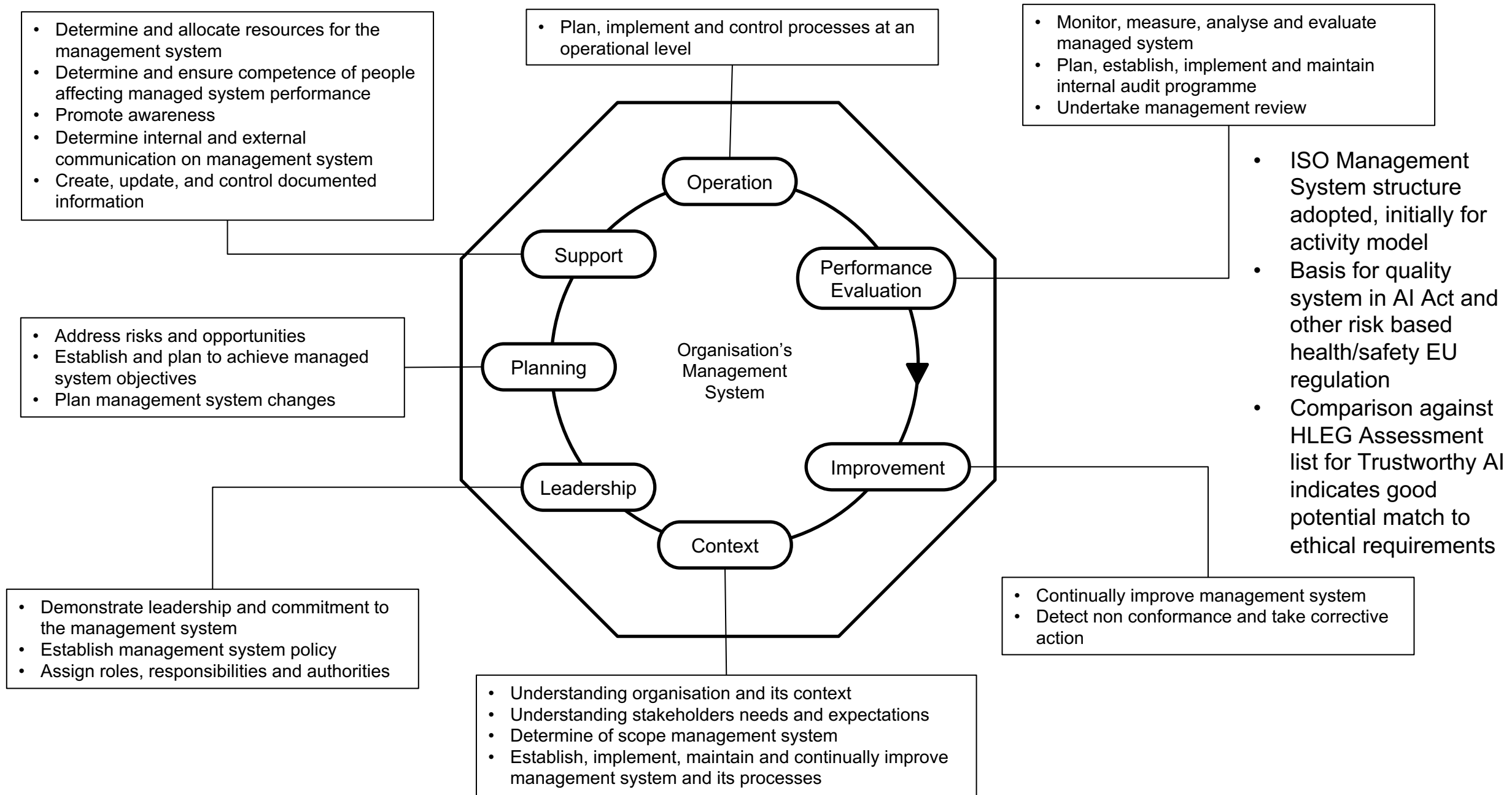


- GDPR was fairly 'self-contained'
- Raft of new digital EU legislation emerging:
 - AI Act,
 - AI Liability
 - Data Governance Act
 - Data Act
 - Digital Service Act
 - Digital Market Act
- Framework relies upon availability of Harmonised Standards for it technical conformance

Legal-to-Standards Mapping: AI Act



ISO Harmonised Structure for Management System Structure



Next Steps: Expanding concept and requirement modelling to AIMS Control and their links to other SC42 standards

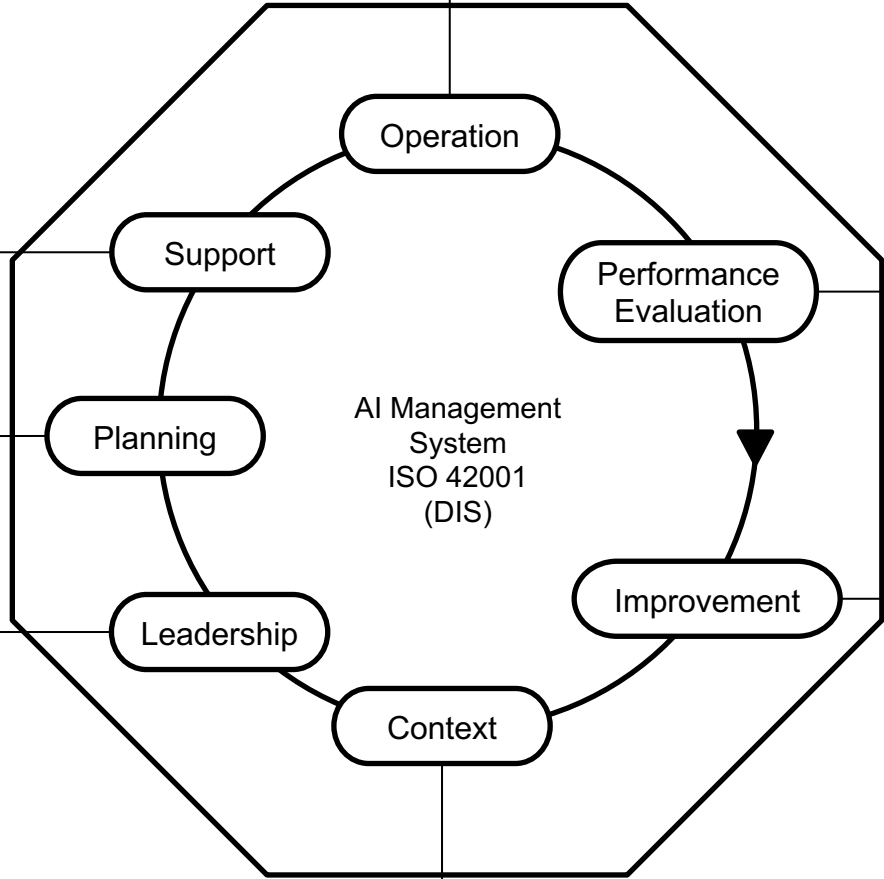
- Resources for AI: data, tooling, system/computing, human (B.4)(B.7)
- Quality of Data Resources ISO/IEC 5259
- Tooling and resources for ML ISO/IEC 23053:

- AI system development lifecycle, responsible AI development (B.6) see ISO/IEC DIS 22989 and ISO/IEC CD 5338 AI lifecycle processes.
- Verification and validation, including bias ISO/IEC DTR 24027 (Bias) and ISO/IEC TR 24029 (assessment of robustness in neural networks)

- Operation and monitoring B.6.2.5
- Trigger retraining (see ISO/IEC 23053) and assessment of classification (ISO/IEC 4213)
- Quality model for AI (ISO/IEC CD 25059)

- Criteria for performing risk assessment and determining acceptable risks - ISO/IEC DIS 38507 (Governance implication of the use of AI)
- AI risk assessment and treatment – ISO/IEC CS 23894 (AI risk management)
- AI impact assessment (B.5)
- Processes and Objectives for trustworthy use (B.9)

- Scope of AI and other related policies (B.2)
- Internal organisation and reporting (B.3) ISO 37002 (whistleblowing management system)
- Third party relationship (B.10)



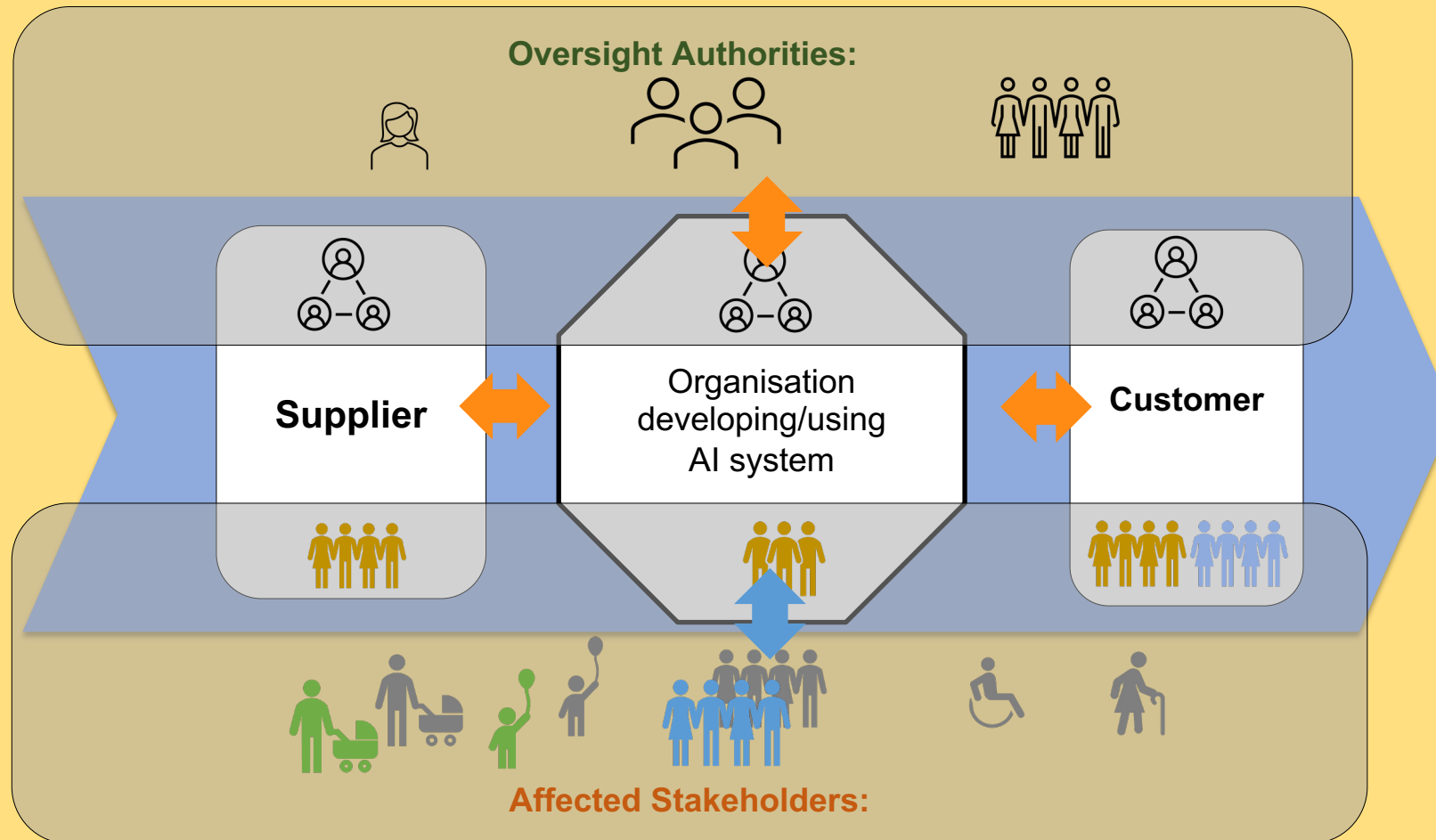
- Current state of controls specified in ISO/IEC AI Management system(AIMS) 42001
- Provisional: 42001 still under development and referenced standard also mostly under development

- Include AI providers, producers, customers, partners
- AI lifecycle and stakeholder ISO/IEC DIS 22989 (Ai terms and concepts)
- PII roles ISO/IEC 29100 (security techniques – privacy framework)

- Information for interested parties (B.8)
- Communication of incidents (B.8.5) ISO/IEC 27001 Security management system requirements, ISO/IEC 27701 privacy information management requirements

Role of Stakeholders in Bridging fundamental rights, ethics and management systems

Societal Context



- Labour Practices (workers)
- The Environment (future generations)
- Fair Operating Procedures (suppliers, customers, regulators)
- Consumer Issues (consumers)
- Community Involvement and Development (local communities)
- Human Rights (everyone)

AI Act: Three little words

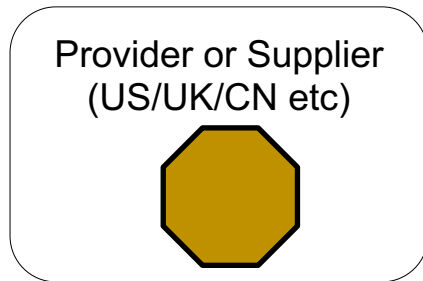
Preamble	Peace – common values	Universal values	Diversity, etc	Rights more visible	Reaffirms const. and int'l rights	Rights, duties, responsibilities	Rights, freedoms and principles
I Dignity (Articles 1–5)	1 Human dignity	2 Life	3 Integrity of the person	4 Torture and inhuman degrading treatment or punishment	5 Slavery and forced labour		
II Freedoms (Articles 6–19)	6 Liberty and security	7 Private and family life	8 Personal data	9 Marry and found family	10 Thought conscience and religion		
	11 Expression and information	12 Assembly and association	13 Arts and sciences	14 Education	15 Choose occupation and engage in work		
	16 Conduct a business	17 Property	18 Asylum	19 Removal, expulsion or extradition			
III Equality (Articles 20–26)	20 Equality before the law	21 Non-discrimination	22 Cultural, religious and linguistic diversity	23 Equality: men and women	24 The child	25 Elderly	26 Integration of persons with disabilities
IV Solidarity (Articles 27–38)	27 Workers right to information and consultation		28 collective bargaining and action	29 Access to placement services	30 Unjustified dismissal	31 Fair and just working conditions	
	32 Prohibition of child labour and protection of young people at work		33 Family and professional life	34 Social security and assistance	35 Health care	36 Access to services of general economic interest	
	37 Environmental protection	38 Consumer protection					
V Citizens' rights (Articles 39–46)	39 Vote and stand as candidate to EP	40 Vote and stand as candidate at municipal elections		41 Good administration	42 Access to documents	43 European ombudsman	
	44 Petition (EP)	45 Movement and residence	46 Diplomatic and consular protection				
VI Justice (Articles 47–50)	47 Effective remedy and fair trial		48 Presumption of innocence and right of defence		49 Legality and proportionality of criminal offences and penalties		50 <i>Ne bis in idem</i>
VII General provisions (Articles 51–54)	51 Application	52 Scope and interpretation	53 Level of protection	54 Prohibition of abuse of rights			

- Reference to European Fundamental Rights
- Large expansion in product certification scope

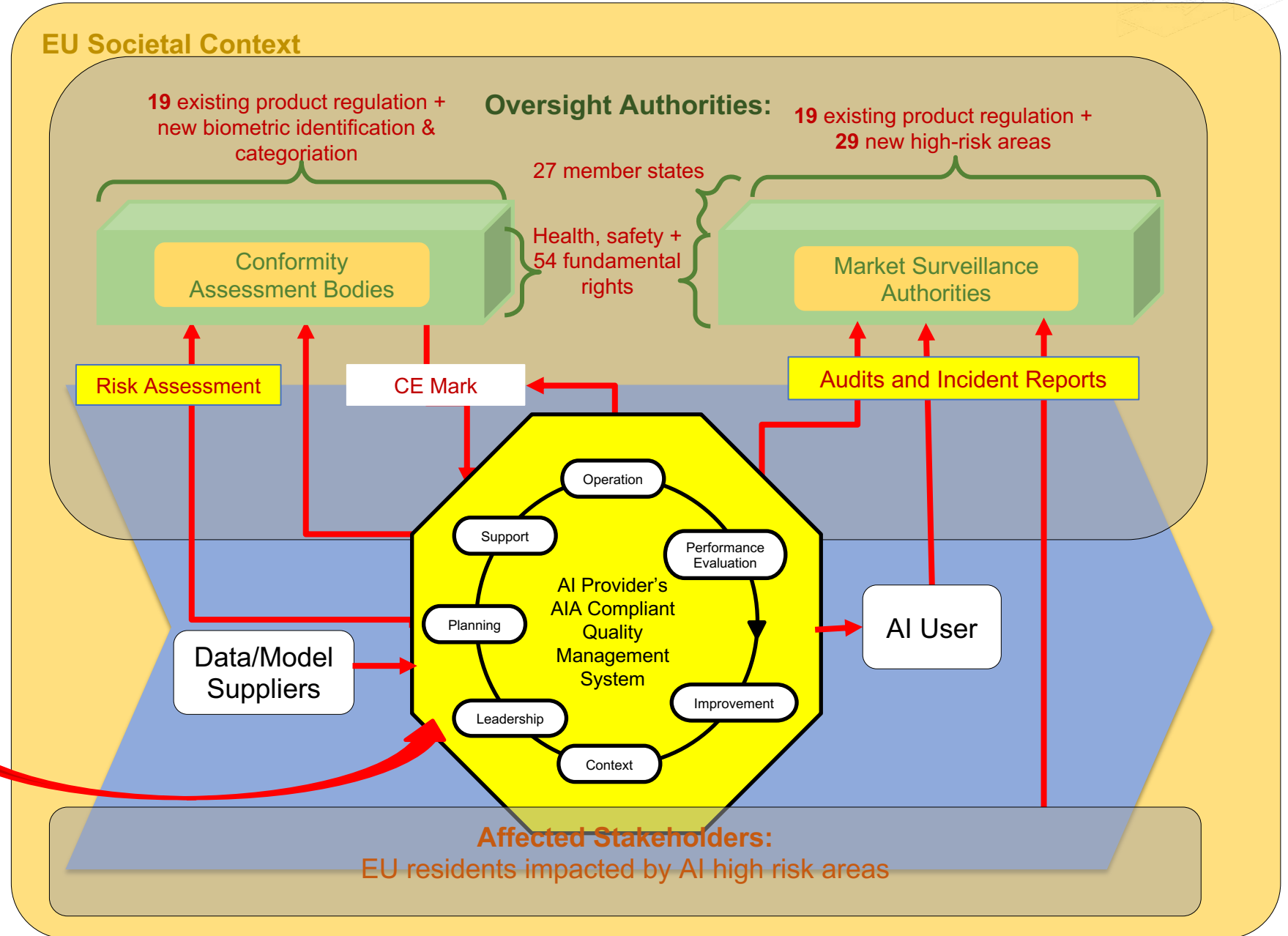
Information Sharing Requirements for the AI Act



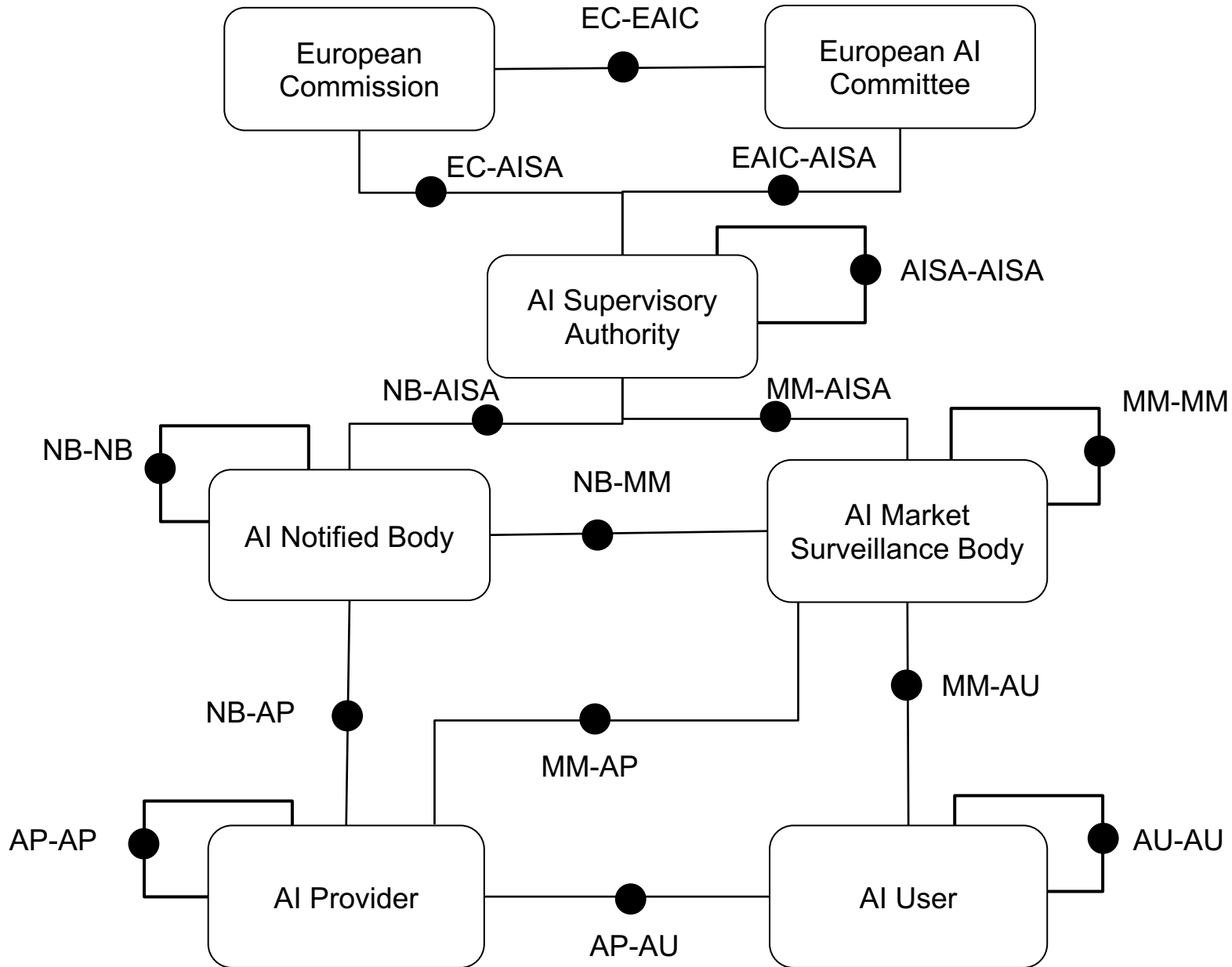
- **Consistent info exchange** between AI providers, Users, Suppliers, Regulators and Affected Stakeholders
- Deploying technology across different application areas
- Mapping from extant proprietary trustworthy AI certification schemes



- Mapping from certification in other Jurisdictions



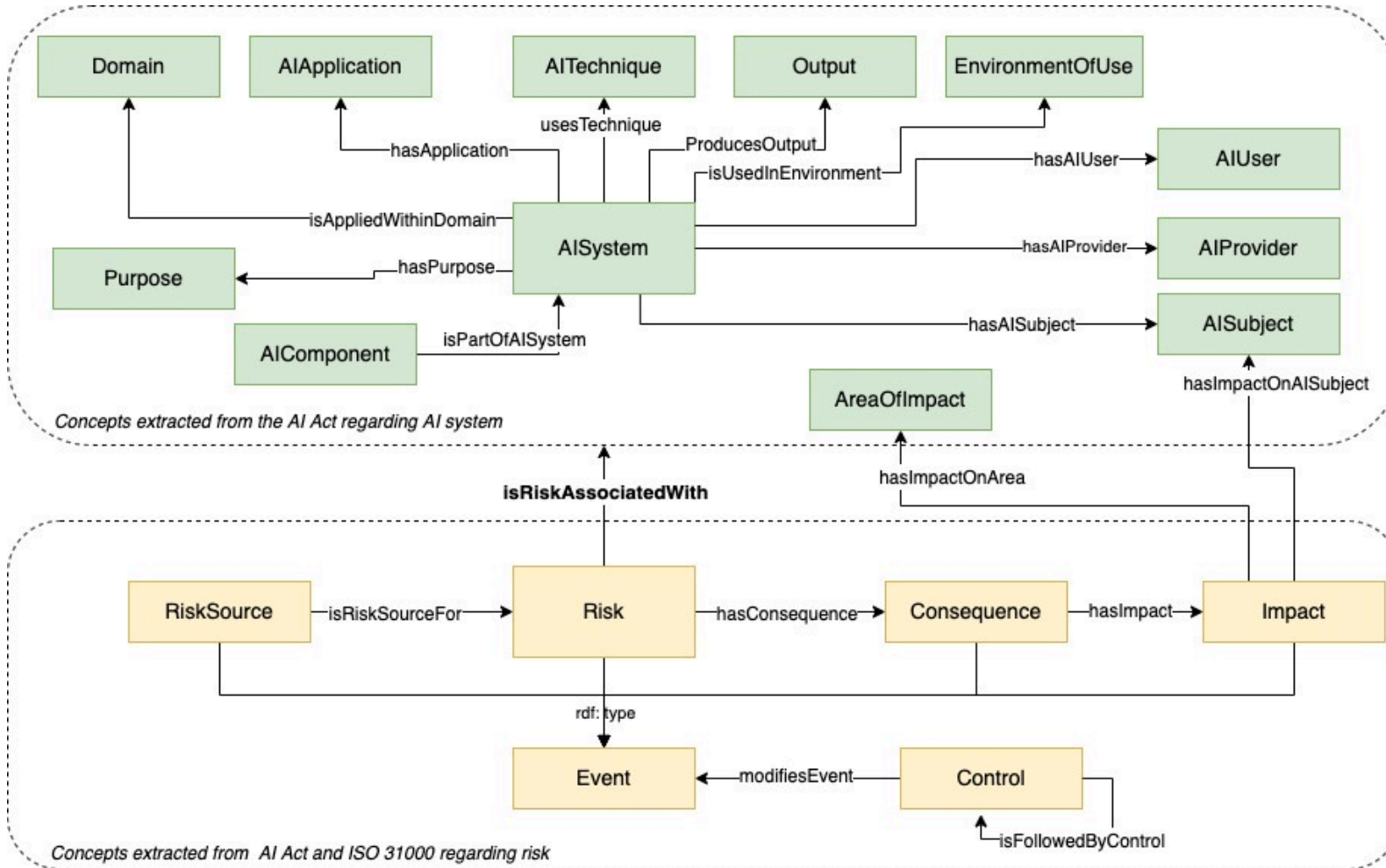
Initial analysis of AI Act to extract actors and exchanged entities



- Initial requirement extracted to model risk activities and entities
- Mapping to ISO/IEC SC42 23894 AI risk management
- Sample of Competence Questions:

R1	What are the risks associated with the AI system?	Art 9(2)(b)	Risk
R2	What are the sources of the risk?	Art 15(4)	RiskSource
R3	What are the consequences of the risk?	ISO 31000	Consequence
R4	What are the harmful impacts of the risk?	Art 5, Art 7	Impact
R5	What measures are in place to prevent/ detect/ mitigate/ eliminate the risk source/ risk/ consequence/ impact?	Art 9(2)(d)	Control
R6	What is the probability of source/ risk/ consequence/ impact happening?	Art 5(2)(b), Art 7(1)(b)	Likelihood
R7	How severe is the harm caused by the system?	Art 5(2)(b), Art 7(1)(b)	Severity
R8	Who is impacted by the AI system?	Art 7(2)	AISubject
R9	What is impacted by the AI system?	Art 7(1)(b)	AreaOfImpact

Application: From term, concepts and requirement mode to operational support models

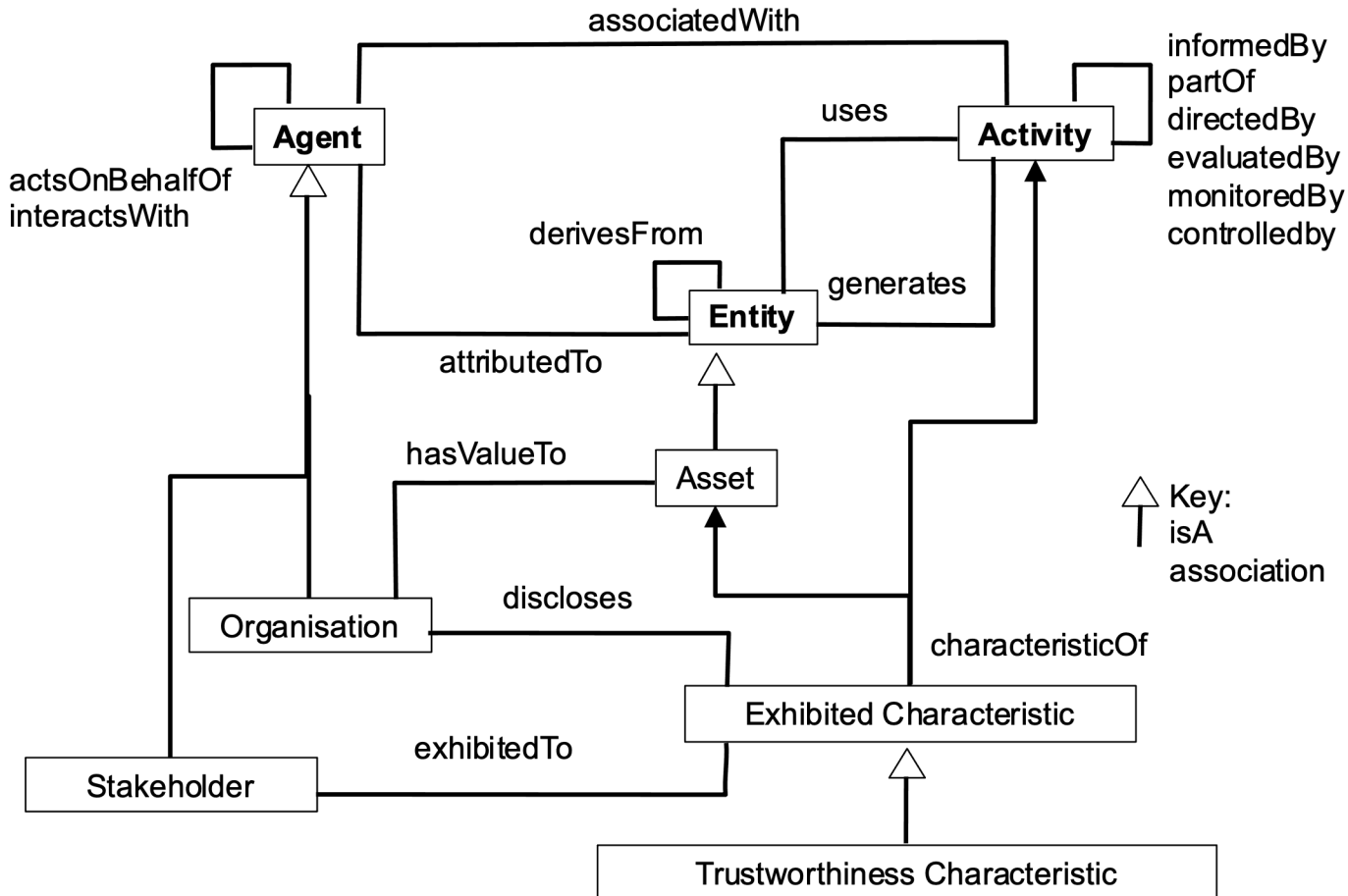


- Bridges risk requirement of AIA with ISO risk management
- Semantic interoperability (with AIA traceability) for document generation, checking and exchange
- E.g. knowledge graph constraints definition for checking high-risk application status

Trustworthy AI Terms and Concepts Catalogue:

- To help actors navigate and reconcile requirements from multiple sources by:
 - Extracting, classifying and mapping terms and concepts from sources in an open knowledge format
 - Provide a platform for navigating terms and concepts in support of specific use cases
- Modelling using Semantic Web Ontologies
 - *Standardised* knowledge graph schema and query/constraint tools (W3C)
 - *Extensible and inter-linkable semantic models* – proven for bridging between knowledge sources
 - *Tool support* for publishing open schema on web and for use operationally in graph databases

High level model for capturing and linking governance terms and concepts

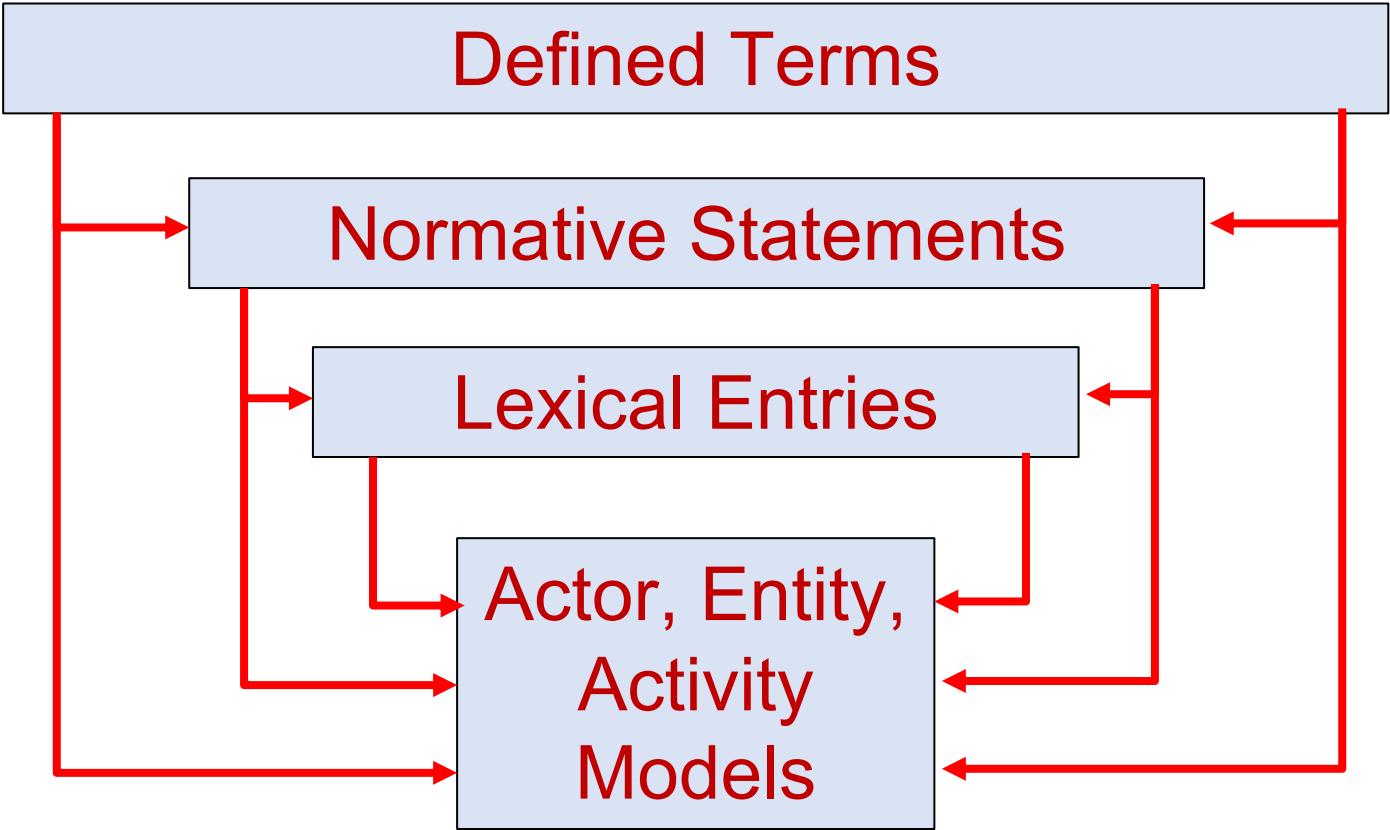
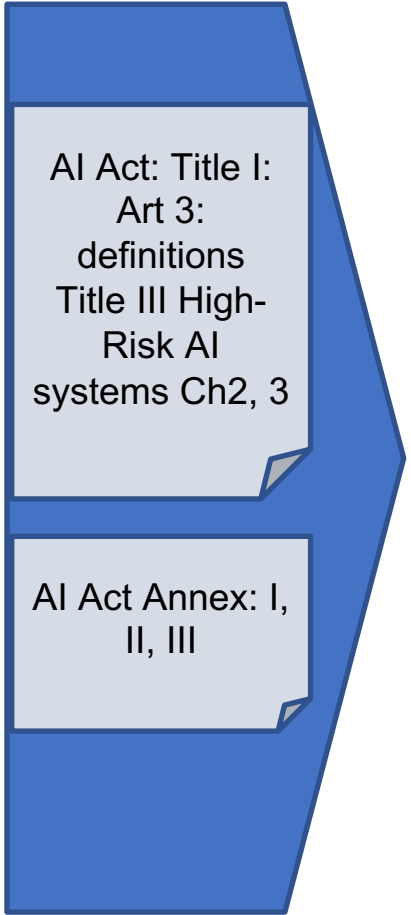


- Goal to extract semantics from source regulations and policies
- Core concepts:
 - Agent: value chain actor, oversight authority or other stakeholder or role with agency in AI/Data governance
 - Activity: value-creating or oversight activities
 - Entity: asset in value chain and related governance artefacts
- Also, use standardized management system structure from ISO to specialize semantic model to governance activities

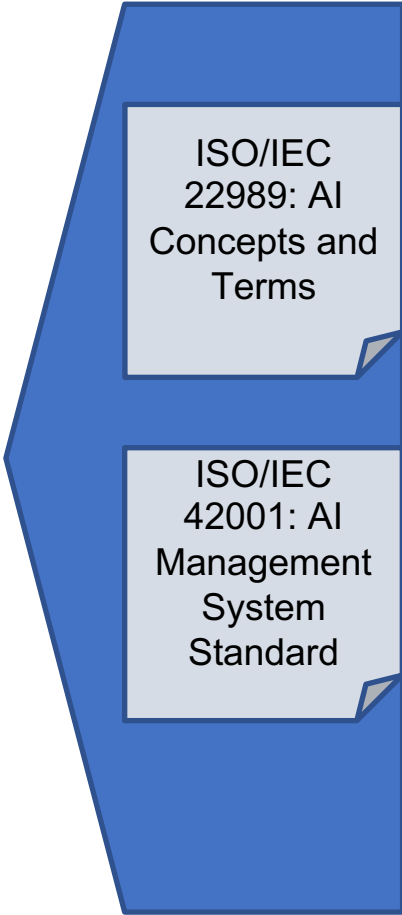
Based on DOI: 10.5772/intechopen.97478

Layered Semantic Modelling of Requirements from separately developed models

Draft AI Act Text



ISO IEC SC42
Standards



Example: Modelling Concepts from ISO/IEC DIC 42001: AI Management System Standard

“3.1
organization
 person or group of people that has its own functions with responsibilities, authorities and relationships to achieve its objectives (3.6)
Note 1 to entry: The concept of organization includes, but is not limited to, sole-trader, company, corporation, firm, enterprise, authority, partnership, charity or institution, or part or combination thereof, whether incorporated or not, public or private.
Note 2 to entry: If the organization is part of a larger entity, the term “organization” refers only to the part of the larger entity that is within the scope of the AI management system (3.4)”

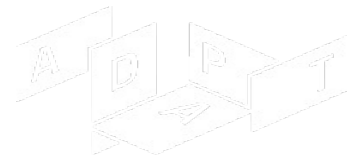
“3.2
interested party (preferred term)
stakeholder (admitted term)
 person or organization (3.1) that can affect, be affected by, or perceive itself to be affected by a decision or activity”

* rdf:type skos:Concept	dct:source	skos:altLabel	skos:prefLabel ""@en	skos:definition ""@en	skos:broader	skos:related	skos:scopeNote	skos:scopeNote
organization	3.1		organization	person or group of people that has its own functions with responsibilities, authorities and relationships to achieve its objectives (3.6)		objective	Note 1 to entry: The concept of organization includes, but is not limited to, sole-trader, company, corporation, firm, enterprise, authority, partnership, charity or institution or part or combination thereof, whether incorporated or not, public or private.	Note 2 to entry: If the organization is part of a larger entity, the term “organization” refers only to the part of the larger entity that is within the scope of the AI management system (3.4).
interestedParty	3.2	stakeholder	interested party	person or organization (3.1) that can affect, be affected by, or perceive itself to be affected by a decision or activity		organization		

SKOS Concept and Collections extracted from Defined Terms in the Draft AI Act

skos:Collection	skos:member
AiDefinitions	AiSystem, Provider, SmallScaleProvider, User, AuthorisedRepresentative, Importer, Distributor, Operator, PlacingOnTheMarket, MakingAvailableOnTheMarket, PuttingIntoService, IntendedPurpose, ReasonablyForeseeableMisuse, SafetyComponent, InstructionsForUse, Recall, Withdrawal, Performance, NotifyingAuthority, ConformityAssessment, ConformityAssessmentBody, NotifiedBody, SubstantialChange, CeMarking, PostMarketMonitoring, MarketSurveillanceAuthority, HarmonisedStandard, CommonSpecification, TrainingData, ValidationData, TestingData, InputData, BiometricData, EmotionRecognitionSystem, BiometricCategorisationSystem, RemoteBiometricIdentificationSystem, RealTimeRemoteBiometricIdentificationSystem, PostRemoteBiometricIdentificationSystem, PubliclyAccessibleSpace, LawEnforcementAuthorityPublic, LawEnforcementAuthorityOther, LawEnforcement, NationalSupervisoryAuthority, NationalCompetentAuthority, SeriousIncidentPersonPropertyEnvironment, SeriousIncidentInfrastructure
AiSystemTechniquesAndApproaches	MachineLearningApproaches, LogicAndKowledgeBasedApproaches, StatisticalApproaches
ListOfUnionHarmonisationLegislation	ListofUnionHarmonisationLegislationBasedOnThe NewLegislativeFramework, ListofOtherUnionHarmonisationLegislation
ListofUnionHarmonisationLegislationBasedOnThe NewLegislativeFramework	MachineryHarmonisedLegislation, ToysHarmonisedLegislation, RecreationalPersonalWatercraftHarmonisedLegislation, LiftsHarmonisedLegislation, ExplosiveGassesHarmonisedLegislation, RadioEquipmentHarmonisedLegislation, PressureEquipmentHarmonisedLegislation, CablewayInstallationHarmonisedLegislation, PersonalProtectiveEquipmentHarmonisedLegislation, BurningGaseousFuelsHarmonisedLegislation, MedicalDevicesHarmonisedLegislation, InVitroDiagnosticMedicalDevicesHarmonisedLegislation
ListofOtherUnionHarmonisationLegislation	CivilAviationHarmonisedLegislation, TwoOrThreeWheelVehiclesAndQuadricyclesHarmonisedLegislation, AgriculturalAndForestryVehiclesHarmonisedLegislation, MarineEquipmentHarmonisedLegislation, RailSystemsHarmonisedLegislation, MotorVehiclesAndTheirTrailersAndComponentsHarmonisedLegislation, CivilAviationSafetyHarmonisedLegislation
HighRiskAreaAiSystems	RemoteBiometricIdAiSystem, CriticalInfrastructureHighRiskArea, SafetyOfCriticalInfrastructureAiSystem, AccessToEducationAndTrainingAiSystem, AssessmentInEducationAndTrainingAiSystem, RecruitmentForEmploymentAiSystem, DecisionsInEmploymentAiSystem, DecisionOnPublicBenefitsAndServicesAiSystem, DecisionOnCreditAssessmentAiSystem, DecisionOnEmergencyServicesAiSystem, OffencesRiskAssessmentAiSystem, EmotionDetectionForLawEnforcementAiSystem, DeepFakeDetectionForLawEnforcementAiSystem, EvidenceReliabilityAssessmentAiSystem, CrimePredictionAiSystem, ProfilingForLawEnforcementAiSystem, CriminalDataAnalysisAiSystem, LieAndEmotionDetectionForImmigrationAiSystem, ImmigrationDocumentCheckAiSystem, ImmigrationApplicationProcessingAiSystem, JudicialAuthorityAssistanceAiSystem
HighRiskAiAreas	BiometricIdHighRiskArea, CriticalInfrastructureHighRiskArea, EducationAndTrainingHighRiskArea, EmploymentRelatedHighRiskArea, AccessToEssentialServicesAndBenefitsHighRiskArea, LawEnforcementHighRiskArea, MigrationAsylumAndBorderControlHighRiskArea, AdministrationOfJusticeAndDemocraticHighRiskArea, BiometricIdHighRiskArea

SKOS Concept and Collections extracted from Defined Terms from ISO/IEC FDIS 22989: AI Concepts and Terms



skos:Collection	skos:member
ArtificialIntelligenceCollection	Agent, AiEngineeredSystem, AiDisipline, AiSystem , Autonomy, ApplicationSpecificIntegratedCircuit, Automatic, BayesianNetwork, CognitiveComputing, ContinuousLearning, ComputerVision, Connectionism, Controllability, DataMining, DecisionTree, ExpertSystem, GeneralAiSystem, GeneticAlgorithm, Heteronomy, HumanMachineLearning, MachineLearningHyperparameter, Inference, InternetOfThings, IotDevice, LifeCycle, Model, NarrowAiSystem, MachineLearningParameter, Performance, AiPlanning, Robot, Robotics, SemanticComputing, SoftComputing, SymbolicAiSystem, SubsymbolicAi, Task
MachineLearningCollection	DataAnnotation, DataQualityChecking, DataAugmentation, DataSampling, Dataset, GroundTruth, Imputation, MachineLearningLabel, MachineLearning, MachineLearningAlgorithm, MachineLearningModel, MachineLearningPrediction, ReinforcementLearning, Retraining, Sample, SemiSupervisedMachineLearning, SupervisedMachineLearning, SupportVectorMachine, TestData, TrainedModel, Training, TrainingData, UnsupervisedMachineLearning, ValidationData
NeuralNetworksCollection	ConvolutionalNeuralNetwork, Convolutional, DeepLearning, ExplodingGradient, FeedForwardNeuralNetwork, LongShortTermMemory, NeuralNetwork, AiNeuron, RecurrentNeuralNetwork
TrustworthinessCollection	Accountable, Accountability, Availability, Bias, Control, Explainability, Predictability, Reliability, Resilience, Risk, Robustness, Safety, Stakeholder, OrganizationalTransparency, SystemTransparency, Trustworthiness, Verification, Validation
NaturalLanguageProcessingCollection	AutomaticSummarization, DialogueManagement, EmotionRecognition, InformationRetrieval, MachineTranslation, NamedEntityRecognition, NaturalLanguage, NaturalLanguageGeneration, NaturalLanguageProcessingSystem, NaturalLanguageProcessingDiscipline, NaturalLanguageUnderstanding, OpticalCharacterRecognition, PartOfSpeechTagging, QuestionAnswering, RelationshipExtraction, SentimentAnalysis, SpeechRecognition, SpeechSynthesis

Example: Modelling Requirements from ISO /IEC DIS 42001

“4.2 Understanding the needs and expectations of interested parties

The organization shall determine:

- the interested parties that are relevant to the AI management system;
- the relevant requirements of these interested parties;
- which of these requirements will be addressed through the AI management system.”

dcterms:definition	rdf:type	dct:source	rm:elaborates	rm:specifiedBy	skos:related
[The organization shall determine] the interested parties that are relevant to the AI management system .	rm:Requirement	4.2.1	ContextOfTheOrganization	UnderstandingTheNeedsAndExpectationsOfInterestedParties	InterestedParty, ManagementSystem
[The organization shall determine] the relevant requirements of these interested parties ;	rm:Requirement	4.2.2	ContextOfTheOrganization	UnderstandingTheNeedsAndExpectationsOfInterestedParties	Requirement, InterestedParty

Future Challenges

- Revising Models as Acts, Standards and Guidance evolves and interdependencies emerge
- Open AI incident report repository for stakeholders to share risk assessments in high-risk areas with legal and standards concept annotation - <https://www.aiaaic.org/>
 - E.g. Risk transparency sandbox – facilitate risk (and test) spec exchange between MNC, SME and professional/sectoral/civil society bodies
- Mapping service from other standards/guides to AIA/SC42 normative+ terminology models
 - e.g. NIST Risk Management, Microsoft Responsible AI standard, IEEE P7001
- Resources for AI conformance training, especially for SME and Public Sectors
- Sharing of open semantic model derived from standards limited by ISO/IEX/CEN/CENELC copyright.
 - EC needs to address legitimacy issues through open access of standards related to implementing AIA protections of fundamental rights – e.g. free, open access

A World
Leading SFI
Research
Centre



Thank You

QA



HOST INSTITUTION



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

HOST INSTITUTION



PARTNER INSTITUTIONS



University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath
Ireland's Global University



MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University



TUS



Maynooth University
National University of Ireland Maynooth



O'É Gaillimh
NUI Galway