

ATC TRUST STANDARD

Версия 1.0

Собствена методология за оценка и изграждане на съответствие с изкуствения
интелект

Агенция за Технологично Съответствие (АТС)

Кингсофт България ЕООД · ЕИК 204978624
гр. Русе, ул. „Мадарски конник“ № 2, ет. 7

Публикувана версия · септември 2026 г.

1. Въведение и цел

ATC Trust Standard е собствената методология на Агенция за Технологично Съответствие (АТС) за оценка на готовността на организация да отговори на изискванията на Регламент (ЕС) 2024/1689 (EU AI Act), надградена върху международно признатата рамка ISO/IEC 42001 за системи за управление на изкуствения интелект.

Стандартът не твърди формална еквивалентност на ISO сертификация. Той е самостоятелна, прозрачно публикувана методология, изградена да отговори на конкретен практически проблем: регламентът задава изисквания, но не дава на малкия и среден бизнес ясен, приложим път за оценка на собствената му готовност. ATC Trust Standard е този път.

Документът е публично достъпен, за да може всеки клиент, партньор или регулаторен орган да провери методологията точка по точка, вместо да разчита на доверие без проверимост.

2. Обхват на приложение

Стандартът се прилага при всеки ангажимент на АТС, включващ:

- Класификация на риска на AI система по смисъла на чл. 3 от AI Act;
- Оценка на готовността на организация в ролята ѝ на доставчик (provider) или внедрител (deployer);
- Изготвяне или поддръжка на техническа документация по Annex IV;
- Оценка на видимост и коректно представяне пред AI системи (GEO/AI-discoverability готовност);
- Оценка на готовността на организацията за agentic AI внедрявания.

Стандартът не заменя правна консултация. Правните заключения по конкретен случай се дават от юристи партньори на АТС, когато е приложимо.

3. Нормативна и методологична основа

ISO/IEC 42001: международен стандарт за системи за управление на изкуствения интелект — структурата за управление на риска, ролите и непрекъснатото подобрене в ATC Trust Standard следва логиката на този стандарт.

EU AI Act (Регламент (ЕС) 2024/1689): източник на конкретните правни задължения — класификация на риска, изисквания за документация, прозрачност и надзор.

Собствени контроли на АТС: области, които горните два документа все още не покриват изчерпателно — agentic AI поведение, GEO/AI-discoverability готовност, и практическо приложение, съобразено с реалностите на българския пазар.

4. Основни принципи

01 · Прозрачност пред клиента

Всяка констатация се обосновава писмено, с ясен източник — норма от регламента, клауза от ISO/IEC 42001, или собствен контрол на ATC, изрично обозначен като такъв.

02 · Проверимост на методологията

Методологията е публикувана в настоящия документ. Клиент или регулатор може да провери всяка стъпка от оценката, вместо да приема резултата на доверие.

03 · Пропорционалност на надзора спрямо риска

Обемът на изискваните действия е пропорционален на нивото на риск на конкретната AI система — високорисковите системи получават пълния обхват на стандарта, минимално рисковите — само базовите му контроли.

04 · Отговорност, независима от осведомеността

Оценката изхожда от правния принцип, че отговорността на организацията като внедрител съществува независимо от вътрешната осведоменост на ръководството за конкретна употреба на AI в организацията. Стандартът затова включва изрична проверка за недекларирана ('shadow AI') употреба, не само за официално одобрени системи.

5. Контролни области

Осемте контролни области на ATC Trust Standard. Първите пет произтичат пряко от AI Act и ISO/IEC 42001; последните три (F, G, H) са собствени контроли на ATC.

A. Управление и отговорност

Идентификация на ролята на организацията (доставчик и/или внедрител) за всяка отделна AI система. Определяне на отговорно лице за AI съответствие в организацията.

B. Класификация на риска

Всяка AI система се класифицира по четирите нива на риск, заложиени в регламента:

Ниво на риск	Описание	Задължения по ATC Trust Standard
Недопустим	Забранени практики по чл. 5 от AI Act — социален скоринг, манипулативни техники и др.	Незабавно прекратяване на употребата; писмено становище до клиента.
Висок	Системи, влияещи на решения за хора — наемане, кредитиране, здраве и др.	Пълна документация по Annex IV, risk register, човешки надзор, FRIA при нужда.
Ограничен	Чатботове и системи с пряко потребителско взаимодействие.	Задължения за прозрачност (чл. 50); маркиране на AI-генерирано съдържание.
Минимален	Вътрешна автоматизация без пряко въздействие върху трети лица.	Базова AI грамотност на персонала; периодичен преглед на употребата.

C. Техническа документация и проследимост

За системи с висок риск — пълен пакет по Annex IV: архитектура, данни за обучение, мерки за управление на риска, доказателства за тестване. Документацията се третира като жив документ, актуализиран при всяка съществена промяна в системата.

D. Човешки надзор

Определяне на конкретен, документиран механизъм за човешки надзор върху решенията на всяка high-risk система — не абстрактно твърдение за надзор, а конкретен процес с отговорно лице.

E. Прозрачност и маркиране на AI съдържание

Проверка на всички точки на контакт с крайния потребител — чатботове, генерирано съдържание — за съответствие с изискванията на чл. 50 от регламента.

F. Agentic AI поведение (собствен контрол на ATC)

Регламентът и ISO/IEC 42001 все още не адресират изчерпателно автономните AI агенти, способни да предприемат действия без пряка човешка намеса на всяка стъпка. ATC Trust Standard въвежда собствена оценка на: обхвата на автономност на агента, границите на действията, които може да предприема без потвърждение, и логирането на действията му за целите на бъдещ одит.

G. GEO / AI-discoverability готовност (собствен контрол на ATC)

Оценка на това доколко организацията е коректно и точно представена пред AI системи (ChatGPT, Gemini, Perplexity и др.) — структурирани данни, entity consistency, техническа достъпност за AI crawler-и. Не е регулаторно изискване, а практически контрол за репутационен риск: AI система, представяща невярна информация за организацията, създава реална вреда, независимо от липсата на пряко правно задължение.

H. Отговорност на внедрителя при неоторизирана употреба (собствен контрол на ATC)

Проверка за недекларирана ('shadow AI') употреба на AI инструменти от персонала, извън официално одобрените системи. Оценката включва препоръка за политика на употреба и обучение на персонала, независимо дали недекларираната употреба вече е констатирана.

6. Процес на прилагане от АТС

Стъпка 1 — Картографиране: идентифициране на всички AI системи, официално одобрени или не, използвани в организацията.

Стъпка 2 — Класификация: прилагане на контролна област Б за всяка идентифицирана система.

Стъпка 3 — Оценка по контролните области: прилагане на приложимите контроли (А–Н) според нивото на риск на всяка система.

Стъпка 4 — Писмен доклад: констатации, приоритизиран план за действие, и конкретни срокове за изпълнение.

Стъпка 5 — Поддръжка (при избор на абонаментна услуга): периодичен преглед за нови системи, регулаторни промени, или ескалация на риска на вече оценена система.

7. Актуализация и версионизиране

ATC Trust Standard се преразглежда при съществена промяна в EU AI Act, ISO/IEC 42001, или установена национална регулаторна практика. Всяка версия е публично достъпна и предходните версии остават архивирани за целите на проследимост.

Текуща версия: 1.0 · Дата на публикуване: септември 2026 г.

Този документ е собственост на Агенция за Технологично Съответствие / Кингсофт България ЕООД. Възпроизвеждане с търговска цел без изрично писмено съгласие е забранено.