IstitutoComprensivo
CEGLIE - MANZONI LUCARELLI
Via D.co Di Venere, 21/23 70129 BARI-CEGLIE
Tel. 0808653000
C.F.: 93423250724
baic846005@istruzione.it baic846005@pec.istruzione.it
www.istitutocomprensivoceglie.edu.it

Criteri per la classificazione del rischio

La classificazione del rischio delle applicazioni di **Intelligenza Artificiale** secondo l'**AI Act** (Regolamento europeo sull'IA, approvato definitivamente nel 2024) si basa su **quattro livelli di rischio**, con obblighi e divieti diversi. Ecco come si classifica un sistema di IA in modo pratico:

1. Rischio Inaccettabile (vietato)

Un sistema è a rischio "inaccettabile" se **mina diritti fondamentali o sicurezza**.
Sono vietati, per esempio:

- Manipolazione cognitivo-comportamentale dannosa.
- Social scoring da parte di autorità pubbliche.
- Identificazione biometrica in tempo reale in spazi pubblici (con rare eccezioni).
- Giocattoli con IA che creano rischi gravi per i minori.

Come valutarlo
Se il sistema può causare **danni fisici/psicologici gravi** o **limitare diritti fondamentali** → quasi certamente è vietato.

2. Alto Rischio (regolamentato)

Un sistema è ad "alto rischio" se è usato in settori delicati e può incidere significativamente sulle persone.

L'AI Act indica **due categorie principali**:

A. Sistemi di IA integrati in prodotti regolamentati
Esempi:

- Dispositivi medici
- Macchinari industriali
- Ascensori
- Veicoli autonomi
- Dispositivi di sicurezza

Qui il rischio deriva dal fatto che l'IA influisce sulla **sicurezza del prodotto fisico**.

B. Sistemi di IA usati in ambiti critici
Esempi:

- Reclutamento e gestione del personale
- **Accesso a istruzione e formazione**

- **Servizi pubblici essenziali**
- Valutazione delle prestazioni dei dipendenti
- Giustizia e forze dell'ordine
- Procedure di asilo e migrazione

Sono sistemi che prendono decisioni che possono **modificare la vita delle persone**.

Requisito:
Serve **gestione del rischio, documentazione, trasparenza, supervisione umana,** registrazione nel database UE, ecc.

3. Rischio Limitato (obblighi di trasparenza)
Sono sistemi che **non incidono direttamente sui diritti,** ma possono comunque ingannare o confondere.

Esempi:
- Chatbot che sembrano persone.
- Sistemi che generano deepfake.
- Sistemi di IA generativa (come ChatGPT) quando non usati in contesti ad alto rischio.

Obblighi:
- l'utente deve sapere che sta interagendo con un'IA;
- se genera contenuti sintetici → devono essere **etichettati**;
- modelli generativi devono seguire regole di sicurezza e mitigazione.

4. Rischio Minimo o Nessun Rischio (libero uso)
Sono applicazioni che non hanno impatti rilevanti sui diritti o la sicurezza.

Esempi:
- Filtri antispam
- Raccomandazioni di film o musica
- IA nei videogiochi
- Strumenti per editing foto o testo non destinati a prendere decisioni critiche

Nessun obbligo specifico, solo norme generali UE.

IstitutoComprensivo
CEGLIE - MANZONI LUCARELLI
Via D.co Di Venere, 21/23 70129 BARI-CEGLIE
Tel. 0808653000
C.F.: 93423250724
baic846005@istruzione.it baic846005@pec.istruzione.it
www.istitutocomprensivoceglie.edu.it

Schema decisionale per classificare il rischio di una applicazione IA

Premessa
Per classificare il rischio di una applicazione di IA è necessario avere presente

- scopo dell'applicazione;
- contesto in cui viene usata;
- tipo di decisioni che prende;
- dati che utilizza.

Per individuare la categoria di rischio è possibile utilizzare il seguente schema decisionale.

1. **Il sistema può creare rischi gravi per la sicurezza o limitare diritti fondamentali?**
   → **Possibile rischio inaccettabile** (verificare lista nell'AI Act).

2. **È usato in un ambito regolamentato dall'AI Act ad alto rischio?**
   - Sanità
   - **Istruzione**
   - Giustizia
   - Forze dell'ordine
   - Lavoro e selezione del personale
   - Servizi pubblici essenziali
   - Prodotti industriali regolamentati

   → **Alto rischio**

3. **Interagisce con persone o crea contenuti sintetici che richiedono trasparenza?**
   → **Rischio limitato**

4. **Non ha impatti critici e non rientra nelle categorie precedenti?**
   → **Rischio minimo**