Användarmanual & transparensinformation – AI-system (SentinelStack)

Denna manual är avsedd att uppfylla kraven enligt AI-förordningen Artikel 13 och 14 genom att förklara hur AI-systemet fungerar, dess syfte, begränsningar, samt hur användaren interagerar med det.

# 1. Vad gör AI-systemet?

- Identifierar potentiellt skadlig eller oetisk kommunikation (t.ex. grooming, phishing)  
- Detekterar deepfakes eller manipulation i innehåll (text, ljud, nätverk)  
- Ger notifieringar eller blockerar misstänkt aktivitet beroende på systeminställningar

# 2. Hur fungerar det?

- AI-modellen analyserar input lokalt (text, bild, ljud eller domäntrafik)  
- Systemet genererar en lokal logg och, vid behov, en varning  
- Varningar skickas till administratör eller vårdnadshavare via Telegram, PDF eller e-post

# 3. Begränsningar

- Systemet är inte 100 % korrekt och kan ge falska positiva eller missa hot  
- Ingen åtgärd sker utan aktivering av användare (local-first design)  
- AI-beslut är stödjande, inte slutgiltiga – användaren avgör åtgärder

# 4. Vad har jag som användare för rättigheter?

- Du har rätt att bli informerad om att AI används  
- Du kan begära att all lokal data raderas  
- Du kan begära utdrag av beslut, loggar eller exporterat PDF-paket

# 5. Hur styr jag AI-systemet?

- Kommandon via CLI: `--enable`, `--disable`, `--export`, `--wipe`  
- Via Telegram: `/status`, `/log`, `/screenshot`, `/shutdown`  
- GuardianAgent: knappar för nödstopp, logg, inställningar

# 6. Kontakta oss

Om du har frågor om detta AI-system, kontakta:  
SentinelStack AB – info@sentinelstack.tech