The European Artificial Intelligence Act

A Comprehensive Framework for AI Governance

Prof. Fedeli Massimo - Tutti i diritti riservati

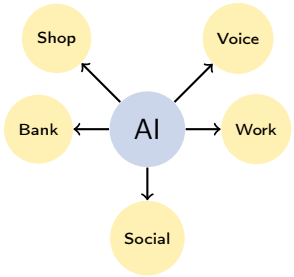
IIS Fermi Sacconi Cpia - Ascoli Piceno

January 5, 2026

The AI Revolution in Our Daily Lives

AI is everywhere:

- Online recommendations and suggestions
- Voice assistants on smartphones
- Credit evaluation systems
- CV screening for job applications
- Social media content filtering



The need for regulation:

- Protect fundamental rights
- Ensure safety and transparency
- Balance innovation with protection

The AI Act: A Historic Achievement

Key Milestones

- **December 9, 2023**: Provisional agreement reached
- **First in the world**: Comprehensive AI regulatory framework
- **Landmark legislation**: Sets global standards for AI governance

What it defines:

- What can and cannot be done with AI
- Guarantees for citizens
- Responsibilities for companies

Who it affects:

- AI developers and providers
- Companies using AI systems
- Public institutions
- Citizens and end users

An Innovative Approach: Risk-Based Regulation

The fundamental principle:

"Higher risk = Stricter rules"

Few obligations

Strict requirements

PROHIBITED

Finding the balance between protection and innovation

Unacceptable Risk: Prohibited AI Practices

These AI applications are completely banned in the EU:

1. **Behavioral manipulation systems**
   - Exploiting vulnerabilities (age, disability, economic status)
   - Subliminal or deceptive techniques
2. **Social scoring systems**
   - Rating citizens based on behavior or characteristics
   - Mass social surveillance by public or private entities
3. **Biometric categorization based on sensitive data**
   - Inferring race, political opinions, sexual orientation
   - Deducing religious or philosophical beliefs

Principle

These practices are incompatible with EU fundamental values and the Charter of Fundamental Rights.

More Prohibited Practices

4. Emotion Recognition

Prohibited in:

- Workplace environments
- Educational institutions

Exception:

- Medical and safety purposes
- Example: monitoring pilot fatigue

More Prohibited Practices

5. Facial Image Scraping

- Untargeted collection from internet
- From CCTV systems
- To create/expand facial recognition databases

More Prohibited Practices

6. Real-Time Facial Recognition

General prohibition for law enforcement in public spaces

Limited exceptions:

- Targeted search for victims
- Prevention of specific threats
- Detection of serious crimes

Requires judicial authorization and strict oversight

Concrete Example: Emotion Recognition Ban

Prohibited Scenario

Company X installs emotion detection cameras

- Monitors employee facial expressions
- Analyzes engagement levels during meetings
- Uses data for performance evaluations
- Claims to detect stress or happiness

Result: PROHIBITED

School Example

University installs AI to detect student attention

- Cameras analyze facial expressions
- System flags distracted students
- Data shared with professors
- Affects participation grades

Result: PROHIBITED

Concrete Example: Social Scoring Systems

Prohibited: CitizenScore System

City government implements AI scoring:

- Tracks social media, shopping, social connections
- Assigns score 0-1000 to each citizen
- High scorers: priority housing, faster permits
- Low scorers: service restrictions, higher rates

Result: ABSOLUTELY PROHIBITED

Why Banned

- Mass surveillance
- Discriminatory access
- Chills [illegible]

Real Context

China's Social Credit:

- Millions affected
- Restricts travel, jobs

Concrete Examples: Biometric Violations

Prohibited: Biometric Categorization

Airport Security System

- AI analyzes facial features
- Infers ethnicity, religion
- Flags profiles for screening
- Appearance-based only

Result: PROHIBITED

Prohibited: Facial Scraping

Tech Company Database

Case Study: Clearview AI

- US company: 3+ billion images
- Sold to law enforcement
- EU countries fined them
- **AI Act explicitly bans this**

The Harm

- Violates privacy massively
- Enables discrimination
- No individual control
- Perpetuates biases
- Chills freedom

High-Risk AI Systems: When AI Needs Special Safeguards

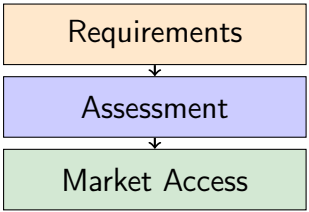
Not prohibited, but requiring strict obligations

Key Requirements

- Conformity assessment before market release
- Quality data for training
- Technical documentation
- Transparency of operation
- Human oversight capability
- Accuracy and robustness
- Cybersecurity measures

Main Principle

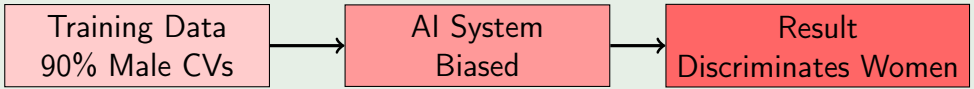
Systems can be used only if they meet quality, safety, and transparency standards



Data Quality: Fighting Discrimination

High-risk systems must be trained with representative datasets

Example (Problem: Biased Recruitment System)



Solution: Representative Training

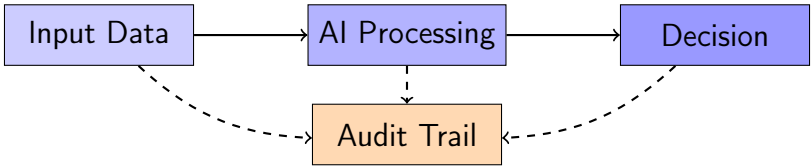
The regulation requires identification and mitigation of discrimination risks

Traceability: Understanding AI Decisions

Why Traceability Matters

High-risk systems must be traceable to reconstruct:

- How the system reached a decision
- Which data were used
- How the system was trained



Complete documentation allows verification and investigation

High-Risk Systems: Practical Categories (1/2)

1. Biometric Identification

- Biometric systems for identifying people
- When not completely prohibited
- Strict oversight required

2. Critical Infrastructure

- Electricity, water, gas networks
- Road traffic management
- Systems whose failure could impact safety

4. Employment

- CV screening systems
- Interview evaluation
- Performance monitoring
- Promotion and dismissal decisions

Common Thread

All these systems can significantly affect people s fundamental rights and safety

High-Risk Systems: Practical Categories (2/2)

5. Essential Services

- Access to public benefits
- Healthcare access
- Social welfare eligibility
- Emergency services dispatch

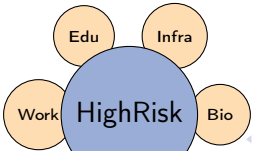
6. Financial Services

- Credit scoring systems
- Creditworthiness evaluation
- Insurance premium calculation
- Loan approval decisions

7. Justice and Law (Italian Request)

- Assisting judges in research
- Interpretation of facts and law
- Alternative dispute resolution

Italy successfully advocated for including judicial AI systems

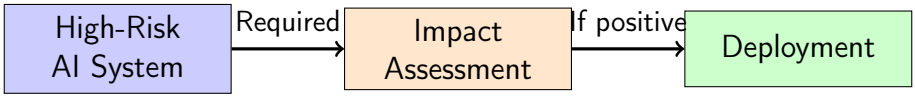


Fundamental Rights Impact Assessment

Additional obligation for public entities and public service providers

The Assessment Must Describe:

1. How the system will be used
2. Who will be affected by it
3. What risks exist
4. What mitigation measures have been adopted



Mandatory for public administration using high-risk AI

Minimal Risk: The Majority of AI Systems

Most AI systems currently used in the EU fall into this category

Examples

- Video games with AI
- Spam filters
- Online shopping recommendations
- Many everyday applications

Voluntary Measures

Encouraged adoption of:

- Codes of conduct
- Best practices
- Self-regulation

Requirements

- No special obligations
- No conformity assessments

Integration

Systems already regulated by:

- Digital Services Act (DSA)

Transparency: The Right to Know

Users must be informed when interacting with AI

1. Chatbots and Virtual Assistants

Clear disclosure required:

- User is interacting with a machine
- Not a real person
- Prevents deception

Example (Good Practice)

"Hello! I'm an AI assistant. How can I help you today?"

2. AI-Generated Content

Must be clearly labeled:

- Texts generated by AI
- Images created by AI
- Audio synthesized by AI
- Videos produced by AI

Machine-Readable Format

- Easily detectable labeling
- Automated verification possible

The Challenge of Deepfakes

What are Deepfakes?

Manipulated videos or audio making it appear that a person said or did things they never actually did

Risks

- Spreading disinformation
- Damaging reputations
- Political manipulation
- Identity fraud
- Erosion of trust

Regulation Requirements

Deepfakes must be:

- Explicitly disclosed as such
- Clearly labeled
- Identifiable by users
- Traceable to source

General Purpose AI and Foundation Models

A new category added during negotiations (Parliament's insistence)

What are General Purpose AI Models?

Systems capable of performing a wide variety of tasks:

- Generate text, images, code
- Translate languages
- Answer questions
- Analyze data

Examples: GPT-4 (OpenAI), Gemini (Google), Claude (Anthropic)

Key Difference

Why Regulate Them?

Models with Systemic Risk

Most powerful models require additional safeguards

Identification Threshold

Models trained with computational power exceeding:

$$10^{25} \text{ FLOPS (Floating Point Operations Per Second)}$$

This threshold can be updated by the AI Office as technology evolves

Systemic Risks

- Mass disinformation
- Coordinated cyberattacks

Required Obligations

- Assess systemic risks
- Mitigate identified risks

Open Source: Balancing Innovation and Safety

Open Source Exemptions

The regulation recognizes the value of:

- Collaborative innovation
- Open-source development
- Research advancement
- Community contributions

Benefits

Free, open-code models can benefit from lighter requirements

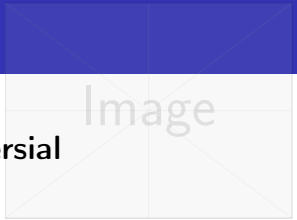
Important Limitation

Even open-source models must comply with stricter safety obligations if they reach the systemic risk threshold

The Balance

- Encourage innovation
- Protect against systemic risks
- Support research community
- Maintain safety standards

The Path to Agreement: Negotiations



The regulation on General Purpose AI was highly controversial

Initial Concerns

Italy, France, Germany worried about:

- Penalizing European companies
- Competitive disadvantage vs. US/China
- Over-regulation stifling innovation
- Compliance costs

The Compromise

- Binding obligations for most powerful models
- Important role for self-regulation through codes of conduct
- Flexibility for innovation
- Strong safety requirements maintained

Parliament's Position

Sanctions: Ensuring Compliance

Strong penalties to ensure the regulation is respected

Sanction Structure

Penalties vary based on violation severity and can be calculated as:

- Fixed amount in millions of euros, OR
- Percentage of worldwide annual turnover

Whichever is higher applies

Most Serious	**Serious**	**Other**
Up to €35M	Up to €15M	Up to €7.5M
or 7% turnover	or 3% turnover	or 1.5% turnover
Prohibited AI	Other requirements	False information

Sanctions: Special Provisions

SMEs and Startups

Special treatment:

- Lower of the two amounts applies
- Recognition that large fines could be devastating
- Proportionality principle

EU Institutions

No exemptions:

- EU agencies subject to fines
- European Data Protection Supervisor can impose sanctions
- Rules apply to public and private equally

Example

For an SME violating data requirements:

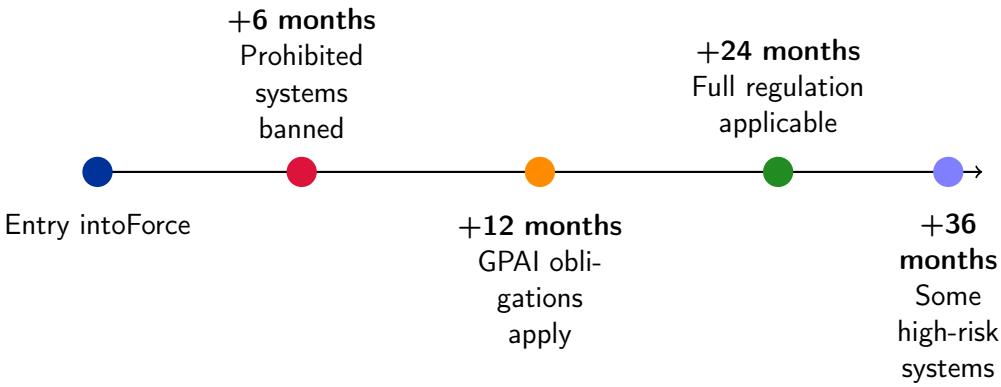
- €35M or 7% turnover
- **Lower amount** is applied

Citizen Rights

- Right to file complaints
- Market surveillance authorities must investigate

When Does It Come Into Effect?

Gradual implementation to allow preparation time



A Model for the World?

The Brussels Effect: EU regulation influences global standards

Precedent: GDPR

The EU's data protection regulation became a de facto global standard:

- Companies adapted globally
- Other countries adopted similar laws
- Set worldwide privacy norms

AI Act Potential

Could follow the same path:

- Tech giants must comply for EU

Different Approaches

United States:

- Market-oriented
- Less interventionist
- Sector-specific rules

China:

- Centralized control
- Social control orientation
- State-driven development

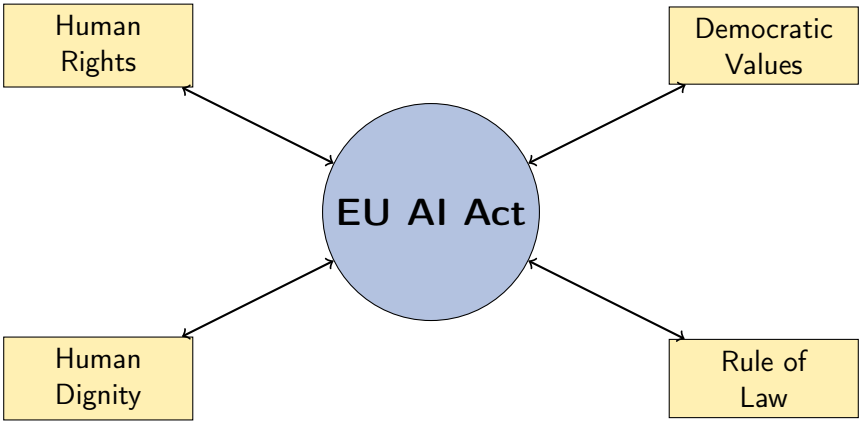
Open Questions and Challenges

The regulation is groundbreaking, but many questions remain

1. **Technological evolution**
   - Will rules remain appropriate in 5-10 years?
   - Can regulation keep pace with AI development?
2. **Balance between protection and innovation**
   - Will Europe fall behind in the global tech race?
   - Can the EU foster AI champions?
3. **Enforcement effectiveness**
   - Will sanctions be sufficient deterrents?
   - Do authorities have adequate resources and expertise?
4. **Global coordination**
   - Will other major economies follow suit?
   - Risk of regulatory fragmentation?
5. **Practical implementation**

Core Principles: The European Choice

The fundamental values guiding the AI Act



Conclusion: A Historic Step Forward

Key Achievements

- First comprehensive AI regulation worldwide
- Risk-based, proportionate approach
- Clear rules on prohibited practices
- Strong safeguards for high-risk systems
- Transparency requirements
- Regulation of powerful AI models
- Meaningful sanctions

Looking Forward

- Implementation will be critical
- Monitoring needed
- Adaptation as technology evolves
- International cooperation

The Challenge

Making the balance work:

AI for Humanity

Thank You!

Questions?

IIS Fermi Sacconi Ceci
Ascoli Piceno, Italy

Understanding AI Regulation
for a Better Digital Future