Open Ethics Transparency Protocol

How self-disclosure can help to bring safer digital diets

Nikita Lukianets
Founder @OpenEthicsAI, PocketConfidant
Fellow, the Alliance of Democracies Foundation
Human decisions are bounded with 4 factors

1. Law
2. Architecture
3. Markets
4. Norms

Lawrence Lessig, 1999. Code and Other Laws of Cyberspace
Disclosure help us to make informed choices
Allowing informed choice

What if we bring labels to digital products?
Motivation

• **Informed consumer choices**: End-users able to make informed choices based on their own ethical preferences and ADM disclosure.

• **Industrial scale monitoring**: Discovery of best and worst practices within market verticals, technology stacks, and product value offerings.

• **Legally-agnostic guidelines**: Suggestions for developers and product-owners, formulated in factual language, which are legally-agnostic and could be easily transformed into product requirements and safeguards.

• **Iterative improvement**: Digital products, specifically, the ones powered by artificial intelligence could receive a nearly real-time feedback on how their performance and ethical posture could be improved to cover security, privacy, diversity, fairness, power-balance, non-discrimination, and other requirements.

• **Labeling and certification**: Mapping to existing and future regulatory initiatives and standards.
Extending The Open Ethics Label

A trust-building platform. Self-disclosure for AI systems.

https://openethics.ai/label
Open Ethics Transparency Protocol

Display “ethical” posture in both human and machine-readable ways.
Open Ethics Transparency Protocol

https://openethics.ai/oetp/
Implementation
Human-in-the-Loop component

AI swarm + Human
Humans confirm

HITL design objective

Personal Data Record

Account ID: Source of the data record: permane
Purpose of the data record processing:
Consent obtained directly
Identifying record

Source

Privacy Policy: https://example.com/privacy

Add comments

Display transparency

This is the preview. After submission of the form, you will receive the code to paste Open Ethics Label on your website.
Progressive Verification with multiple Auditors

1. Disclosure Request
2. OETP Disclosure File
3. Disclosure Validation and Verification Request

Progressive Disclosure Verification

4. Retrieve and Compare Disclosure Integrity
5. Disclosure Verification Request
6. Verification Proof 1
7. Disclosure Verification Request
8. Verification Proof N

9. Verification response
10. Service Request
11. Service Response
Integrity

Open Ethics has generated the following SHA3-512 hash for the snapshot of submitted data.

`e93ba1c91413496d71c839105f0c2c882d3d34d6ed7456103cdd84e0ac4aa6ac6ca35141865cd80c99bb2aad12c0ff306325dd8b555e2fe9c79d7f0b2752cc`

Open Ethics Transparency Protocol

Save the code below as a text file named `oetp.json` in the root of your product’s website. It will allow integrity validation, as well as machine readable scenarios for AI transparency management.

```json
{
    "schema": {
        "name": "Open Ethics Transparency Protocol",
        "version": "0.9.3 RFC",
        "integrity": "e93ba1c91413496d71c839105f0c2c882d3d34d6ed7456103cdd84e0ac4aa6ac6ca35141865cd80c99bb2aad12c0ff306325dd8b555e2fe9c79d7f0b2752cc",
    },
    "snapshot": {
        "product": {
            "url": "letftest.com",
            "description": ""
        }
    }
}
```

Download `oetp.json`
Implications

AI alignment
Safety  
Failure Transparency  
Judicial Transparency  
Responsibility  

Value Alignment  
Personal Privacy  
Liberty and Privacy  
Shared Benefit  

Human Control  
Non-subversion  
Compassion  
Shared Prosperity  

Emotional Value  
Cross-national Security  
Non-influence  
Traceability of decisions
Alternative 1

- Safety
- Failure Transparency
- Judicial Transparency
- Responsibility

- Value Alignment
- Personal Privacy
- Liberty and Privacy
- Shared Benefit

- Human Control
- Non-subversion
- Compassion
- Shared Prosperity

- Emotional Value
- Cross-national Security
- Non-influence
- Traceability of decisions
Alternative 2

- Safety
- Failure Transparency
- Judicial Transparency
- Responsibility

- Value Alignment
- Personal Privacy
- Liberty and Privacy
- Shared Benefit

- Human Control
- Non-subversion
- Compassion
- Shared Prosperity

- Emotional Value
- Cross-national Security
- Non-influence
- Traceability of decisions
Alternative 3

- Safety
- Failure Transparency
- Judicial Transparency
- Responsibility

- Value Alignment
- Personal Privacy
- Liberty and Privacy
- Shared Benefit

- Human Control
- Non-subversion
- Compassion
- Shared Prosperity

- Emotional Value
- Cross-national Security
- Non-influence
- Traceability of decisions
Aligned choice
Use-case

e.example
Public Surveillance Transparency Project (PST)

Citizen in the public space → Scan QR code on the sticker → Read the disclosure → Provide feedback

Station Entrance
1. Visual transparency labels for citizens
2. Machine-readable disclosure
3. Disclosure as Open Data [governance/market]
Market creation: Location-based services

Open Data
McKinsey reference: 5.2 bn EUR
ODI estimates: 1.2 bn EUR

LBS
in Europe alone estimates (2025): 10 bn EUR

https://theodi.org/article/the-value-of-open-data-for-the-private-sector/
https://www.cio.com/article/3626390/6-ways-to-harness-the-value-of-open-data.html
Sharing is caring

Open Ethics Canvas
Talk #ethics, include your #tech team

Download and use it for free

https://openethics.ai/canvas/
Nikita Lukianets
n.lukianets@openethics.ai