

Crypto's problems today  
dAI/dt problems tomorrow

# **Bitcoin/cryptocurrencies are:**

- money out of thin air
- national security threat
- terrorists' best friends
- money laundering facility
- guns purchase Disneyland

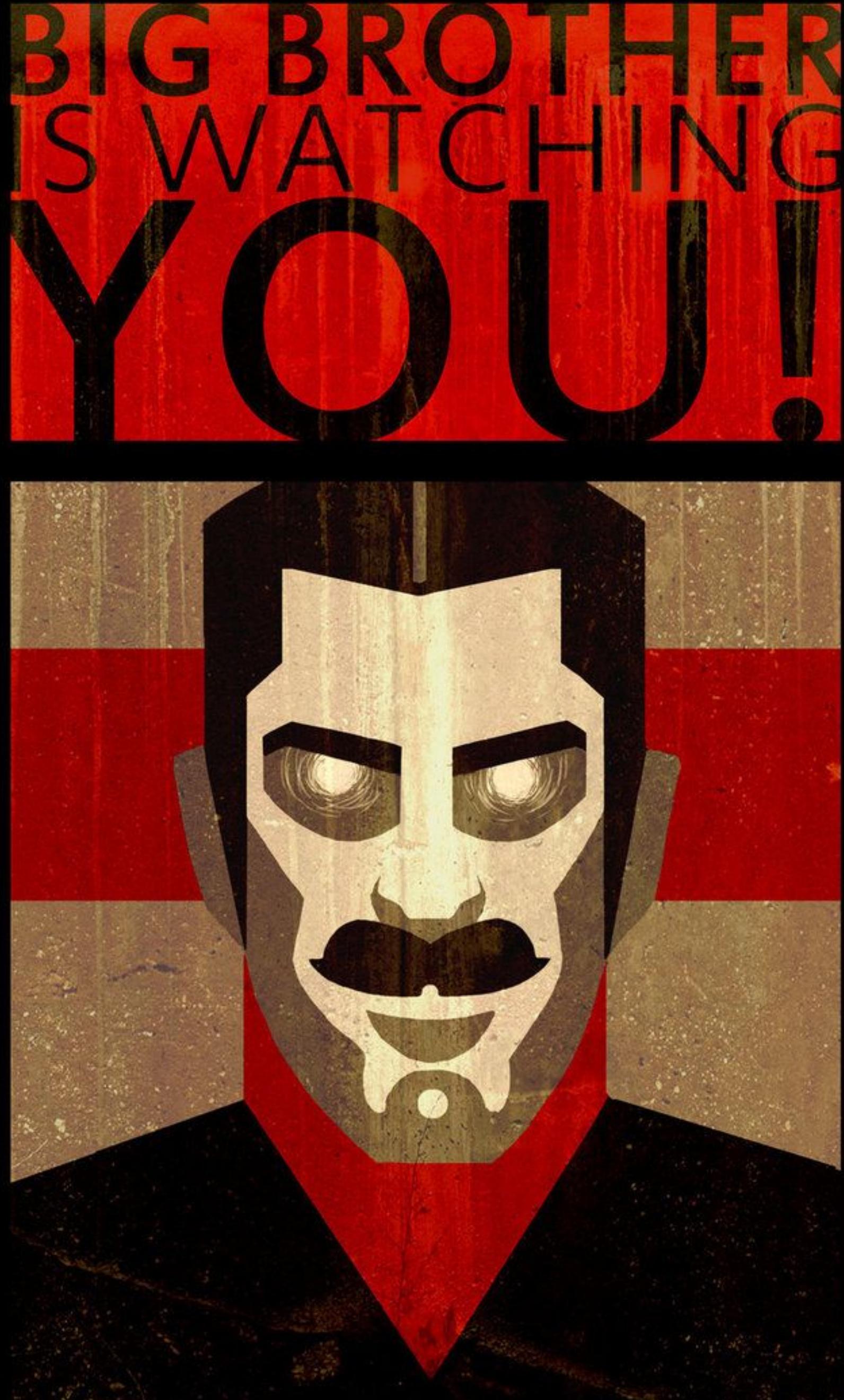
# Bitcoin/cryptocurrencies are:

- guns purchase Disneyland
- cause of your neighbor's drug overdose
- child porn lead investor
- 50 shades of colored people genocide machine
- penguins assassin working for the Global Warming

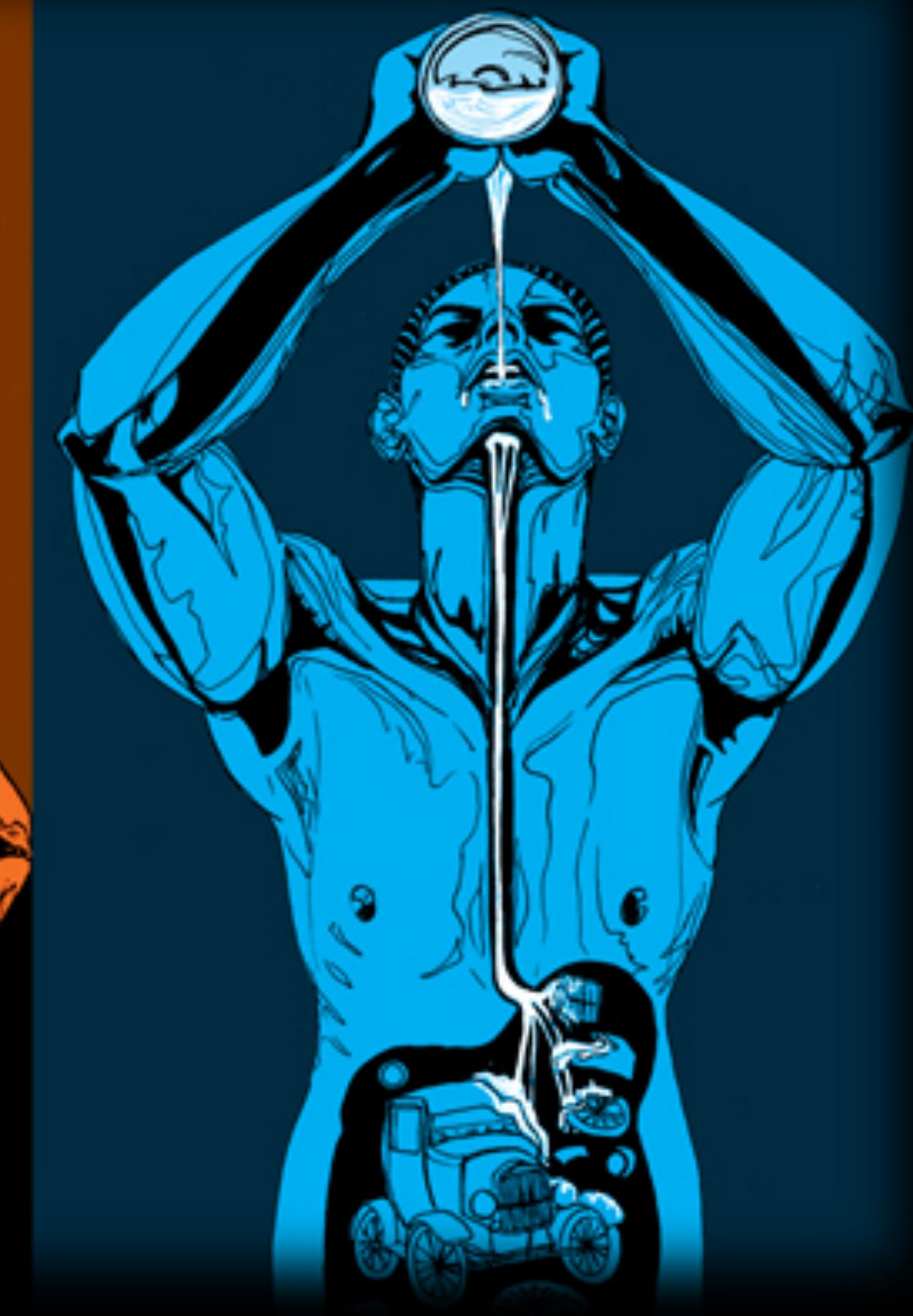


# Orwellian world of today

- “War is peace. Freedom is slavery. Ignorance is strength.”
- “The whole climate of thought will be different. There will be no thought.”
- “Principles: Newspeak, double think, the mutability of the past.”
- “Who controls the past, controls the future: who controls the present controls the past.”
- “No emotion was pure, because everything was mixed up with fear and hatred.”

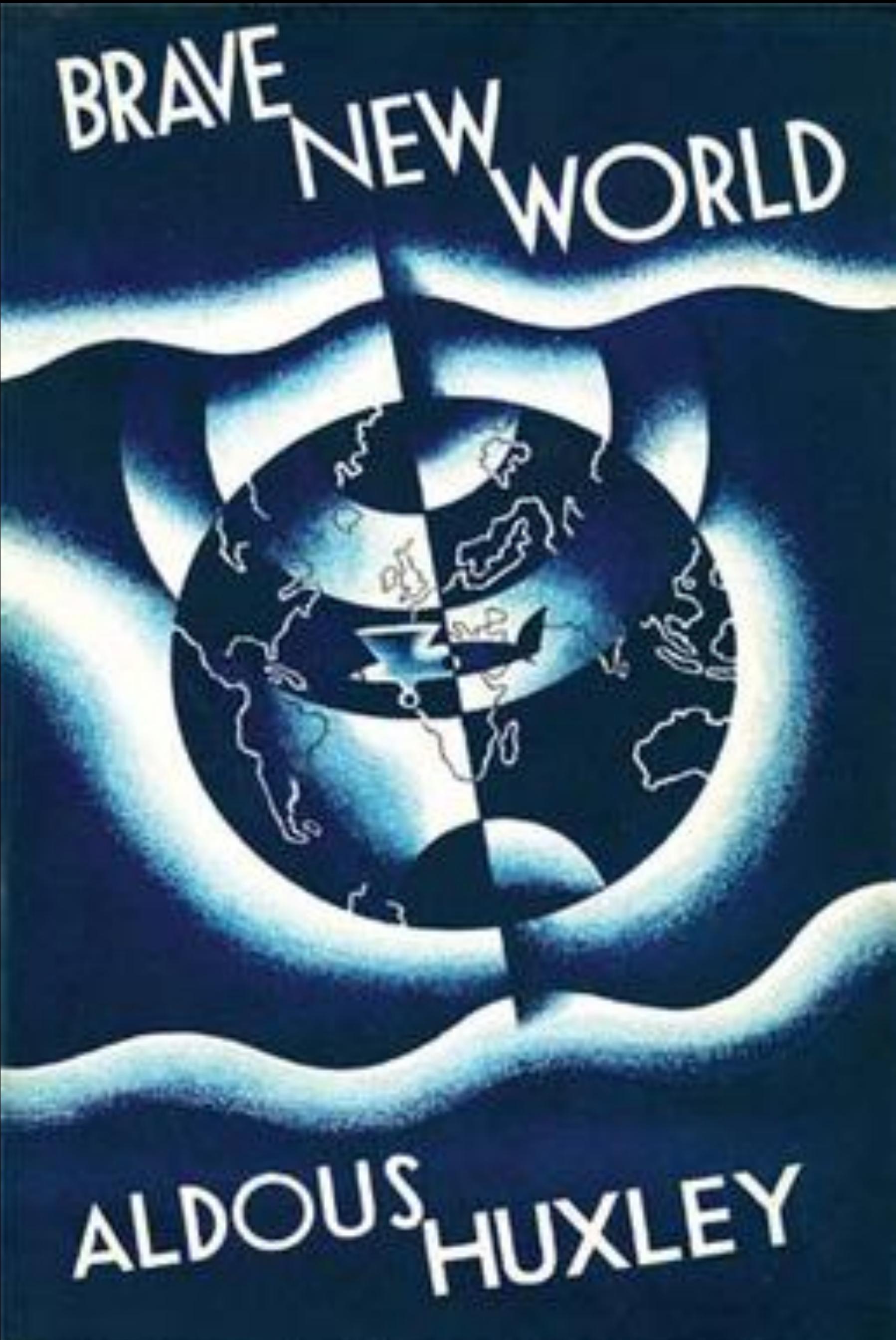


**Is it the only dystopian realm  
we are living in?**



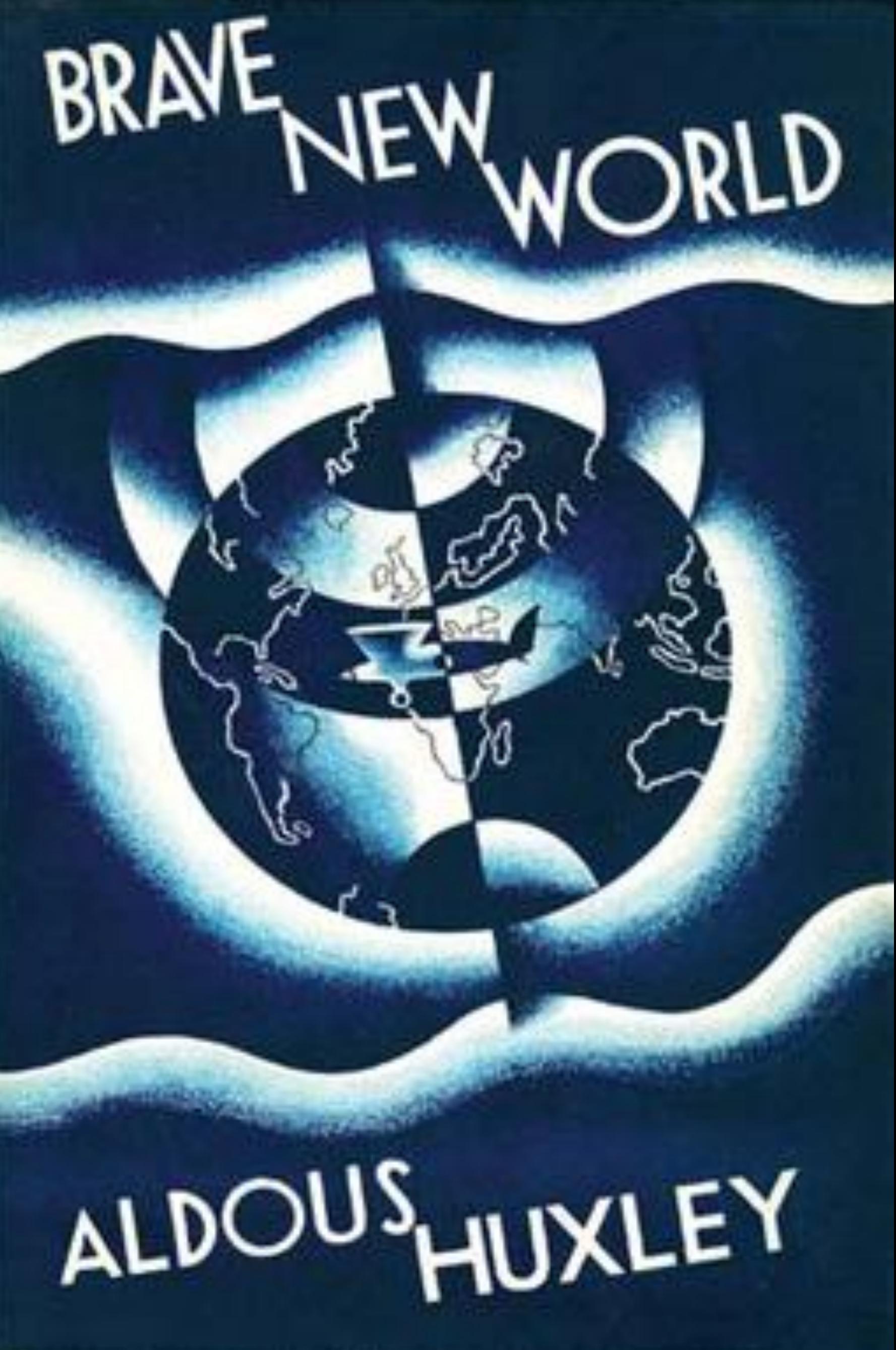
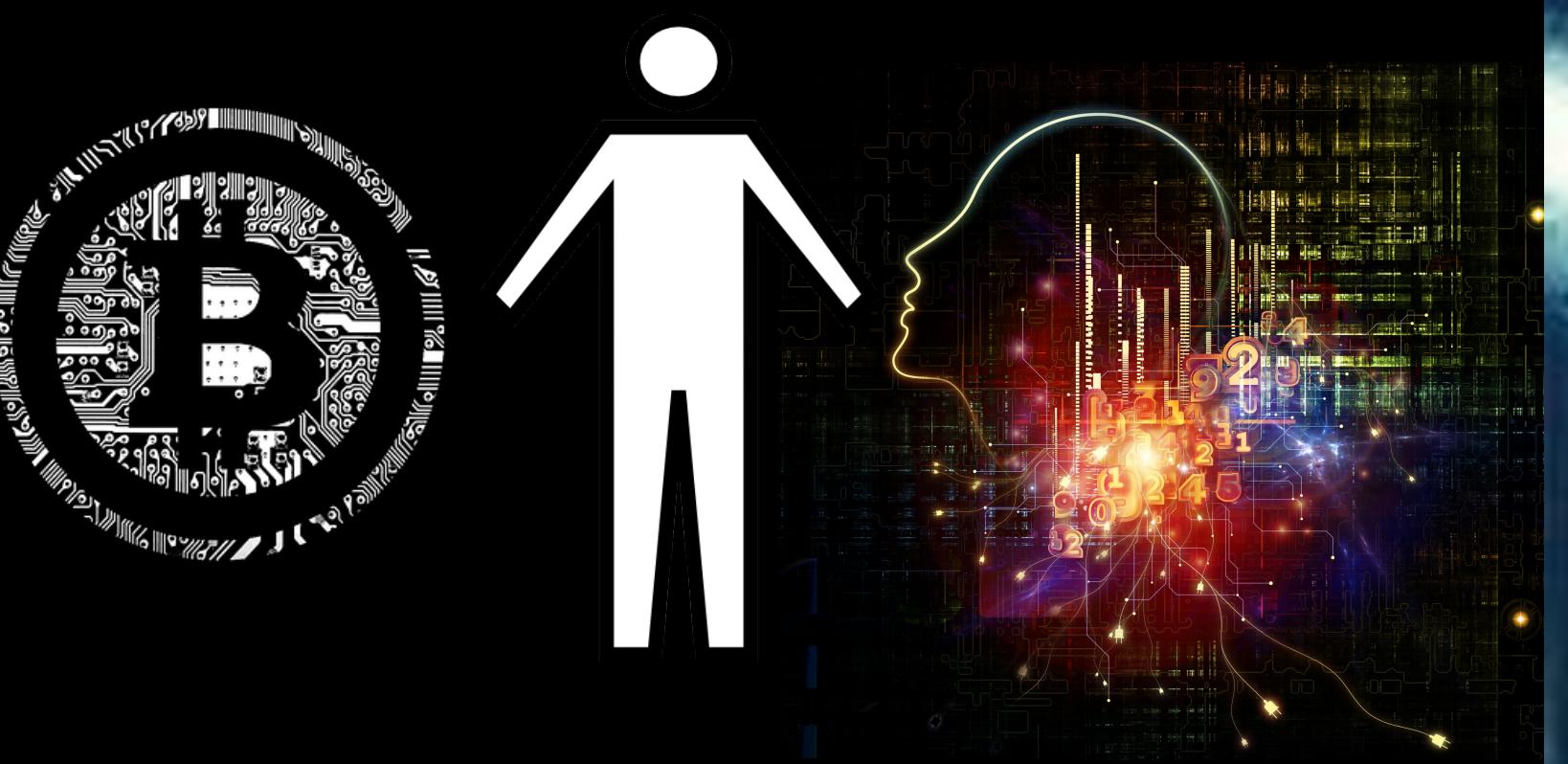
# Huxley's world of today

- "Science is used only insofar as it is a tool for control."
- "Moral education, which ought never, in any circumstances, to be rational."
- "...most men and women will grow up to love their servitude and will never dream of revolution."
- "Most human beings have an almost infinite capacity for taking things for granted."
- "We are not our own any more than what we possess is our own. We did not make ourselves, we cannot be supreme over ourselves. We are not our own masters."
- Soma for a pain-free society.

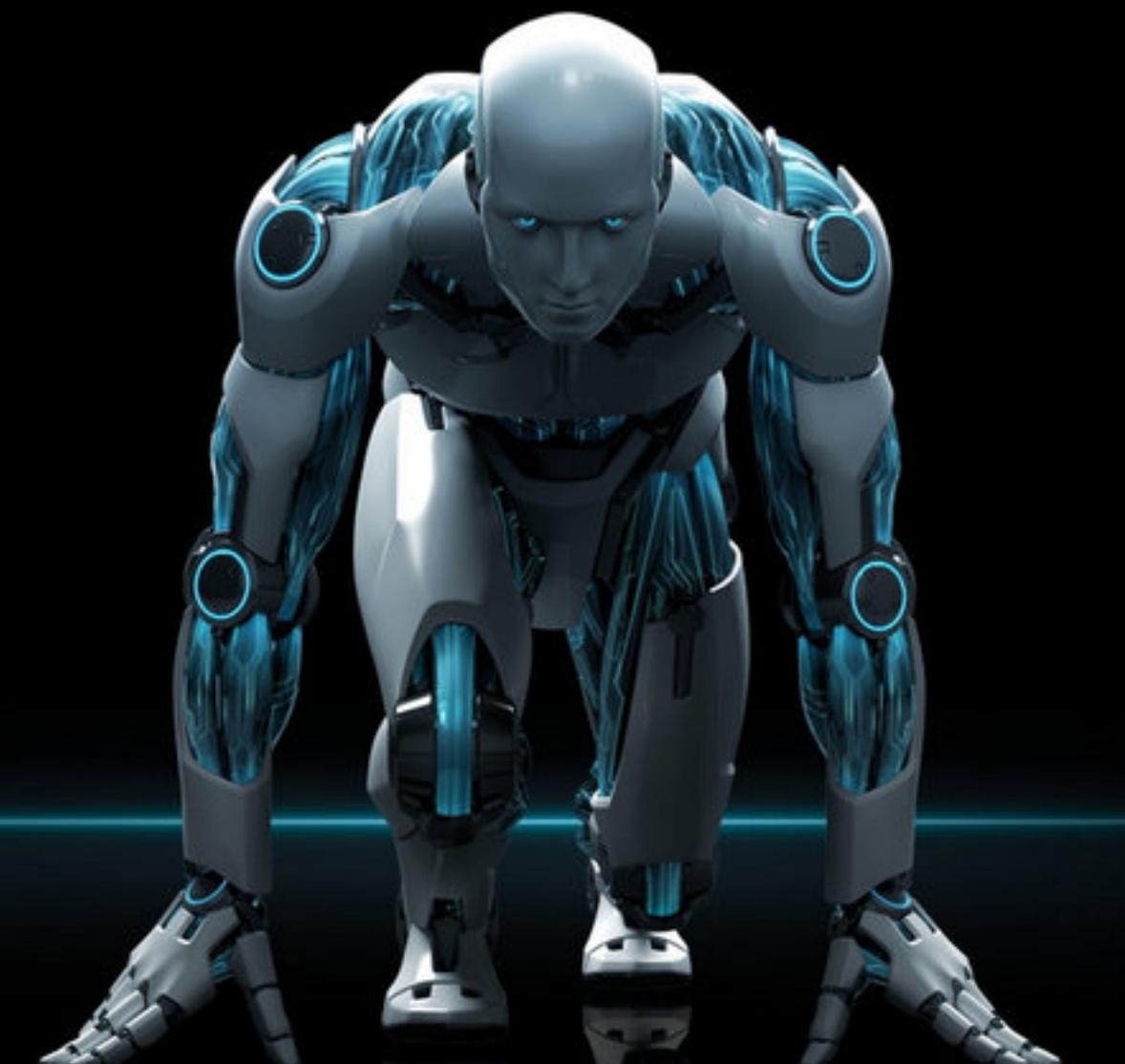


“The world’s stable now. People are happy; they get what they want, and they never want what they can’t get. They’re well off; they’re safe; they’re never ill; they’re not afraid of death; they’re blissfully ignorant of passion and old age; they’re plagued with no mothers or fathers; they’ve got no wives, or children, or lovers to feel strongly about; they’re so conditioned that they practically can’t help behaving as they ought to behave. And if anything should go wrong, there’s soma. Which you go and chuck out of the window in the name of liberty, Mr. Savage. **Liberty!**”

BIG BROTHER  
IS WATCHING  
YOU!



AI will be the next target



# AI

*Today*

- Frees the time for creativity
- Unlocks new level of tech progress
- Helps finding new drugs, treatments, design gene therapy

*Tomorrow*

- Extending own intelligence
- Curing 'uncurable' diseases
- Extending life beyond biological body
- Post-humanity

# AI Regulations

*Today*

- GDPR: pseudo-privacy for the sake of political means
- Bans for autonomous weapon research

*Tomorrow*

- New AI regulations for «keeping the jobs»
- New taxations for AI businesses
- New regulations for life longevity research & usage
- Ban to all DIY activities
- Monopoly for bio data & NN architectures

# Proofs

# Beijing AI principles\*

- Do good: promote the sustainable development of nature and society, to benefit all humankind and the environment, and to enhance the well-being of society and ecology
- Be Diverse and Inclusive: The development of AI should reflect diversity and inclusiveness, and be designed to benefit as many people as possible, especially those who would otherwise be easily neglected or underrepresented in AI applications
- Education and training: Stakeholders of AI systems should be able to receive education and training to help them adapt to the impact of AI development in psychological, emotional and technical aspects



## A. INTRODUCTION

In its Communication of 25 April 2018 and 7 December 2018, the European Commission set out its vision for artificial intelligence (AI), which supports "ethical, secure and cutting-edge AI made in Europe".<sup>5</sup> Three pillars underpin the Commission's vision: (i) increasing public and private investments in AI to boost its uptake, (ii) preparing for socio-economic changes, and (iii) ensuring an appropriate ethical and legal framework to strengthen European values.

tech, independent, right

To support the implementation of this vision, the Commission established the High-Level Expert Group on Artificial Intelligence (AI HLEG), an independent group mandated with the drafting of two deliverables: (1) AI Ethics Guidelines and (2) Policy and Investment Recommendations.

This document contains the AI Ethics Guidelines, which have been revised following further deliberation by our Group in light of feedback received from the public consultation on the draft published on 18 December 2018. It builds on the work of the European Group on Ethics in Science and New Technologies<sup>6</sup> and takes inspiration from other similar efforts.<sup>7</sup>

Over the past months, the 52 of us met, discussed and interacted, committed to the European motto: united in diversity. We believe that AI has the potential to significantly transform society. AI is not an end in itself, but rather a promising means to increase human flourishing, thereby enhancing individual and societal well-being and the common good, as well as bringing progress and innovation. In particular, AI systems can help to facilitate the achievement of the UN's Sustainable Development Goals, such as promoting gender balance and tackling climate change, rationalising our use of natural resources, enhancing our health, mobility and production processes, and supporting how we monitor progress against sustainability and social cohesion indicators. **WTF?**

To do this, AI systems<sup>8</sup> need to be **human-centric**, resting on a commitment to their use in the service of humanity and the common good, with the goal of improving human welfare and freedom. While offering great opportunities, AI systems also give rise to certain risks that must be handled appropriately and proportionately. We now have an important window of opportunity to shape their development. We want to ensure that we can trust the socio-technical environments in which they are embedded. We also want producers of AI systems to get a competitive advantage by embedding Trustworthy AI in their products and services. This entails seeking to **maximise the benefits of AI systems** while at the same time **preventing and minimising their risks**.

In a context of rapid technological change, we believe it is essential that trust remains the bedrock of societies, communities, economies and sustainable development. We therefore identify **Trustworthy AI as our foundational ambition**, since human beings and communities will only be able to have confidence in the technology's development and its applications when a clear and comprehensive framework for achieving its trustworthiness is in place.

This is the path that we believe Europe should follow to become the home and leader of cutting-edge and ethical technology. It is through Trustworthy AI that we, as European citizens, will seek to reap its benefits in a way that is aligned with our foundational values of respect for human rights, democracy and the rule of law.

### Trustworthy AI

Trustworthiness is a prerequisite for people and societies to develop, deploy and use AI systems. Without AI systems – and the human beings behind them – being demonstrably worthy of trust, unwanted consequences may ensue and their uptake might be hindered, preventing the realisation of the potentially vast social and economic

<sup>5</sup> COM(2018)237 and COM(2018)795. Note that the term "made in Europe" is used throughout the Commission's communication. The scope of these Guidelines however aims to encompass not only those AI systems made in Europe, but also those developed elsewhere and deployed or used in Europe. Throughout this document, we hence aim to promote trustworthy AI "for" Europe.

<sup>6</sup> The European Group on Ethics in Science and New Technologies (EGE) is an advisory group of the Commission.

<sup>7</sup> See Section 3.3 of COM(2018)237.

<sup>8</sup> The Glossary at the end of this document provides a definition of AI systems for the purpose of this document. This definition is further elaborated on in a dedicated document prepared by the AI HLEG that accompanies these Guidelines, titled "A definition of AI: Main capabilities and scientific disciplines".

raise the same ethical concerns as AI systems proposing critical medical treatments. Likewise, different opportunities and challenges arise from AI systems used in the context of business-to-consumer, business-to-business, employer-to-employee and public-to-citizen relationships, or more generally, in different sectors or use cases. Given the context-specificity of AI systems, the implementation of these Guidelines needs to be adapted to the particular AI-application. Moreover, the necessity of an additional sectorial approach, to complement the more general horizontal framework proposed in this document, should be explored.

To gain a better understanding of how this guidance can be implemented at a horizontal level, and of those matters that require a sectorial approach, we invite all stakeholders to pilot the Trustworthy AI assessment list (Chapter III) that operationalises this framework and to provide us feedback. Based on the feedback gathered through this piloting phase, we will revise the assessment list of these Guidelines by early 2020. The piloting phase will be launched by the summer of 2019 and last until the end of the year. All interested stakeholders will be able to participate by indicating their interest through the European AI Alliance.

## B. A FRAMEWORK FOR TRUSTWORTHY AI

These Guidelines articulate a framework for achieving Trustworthy AI based on fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union (EU Charter), and in relevant international human rights law. Below, we briefly touch upon Trustworthy AI's three components.

### Lawful AI

AI systems do not operate in a lawless world. A number of legally binding rules at European, national and international level already apply or are relevant to the development, deployment and use of AI systems today. Legal sources include, but are not limited to: EU primary law (the Treaties of the European Union and its Charter of Fundamental Rights), EU secondary law (such as the General Data Protection Regulation, the Product Liability Directive, the Regulation on the Free Flow of Non-Personal Data, anti-discrimination Directives, consumer law and Safety and Health at Work Directives), the UN Human Rights treaties and the Council of Europe conventions (such as the European Convention on Human Rights), and numerous EU Member State laws. Besides horizontally applicable rules, various domain-specific rules exist that apply to particular AI applications (such as for instance the Medical Device Regulation in the healthcare sector).

The law provides both positive and negative obligations, which means that it should not only be interpreted with reference to what **cannot** be done, but also with reference to what **should** be done and what **may** be done. The law not only prohibits certain actions but also enables others. In this regard, it can be noted that the EU Charter contains articles on the 'freedom to conduct a business' and the 'freedom of the arts and sciences', alongside articles addressing areas that we are more familiar with when looking to ensure AI's trustworthiness, such as for instance data protection and non-discrimination.

The Guidelines do not explicitly deal with the first component of Trustworthy AI (lawful AI), but instead aim to offer guidance on fostering and securing the second and third components (ethical and robust AI). While the two latter are to a certain extent often already reflected in existing laws, their full realisation may go beyond existing legal obligations.

Nothing in this document shall be construed or interpreted as providing legal advice or guidance concerning how compliance with any applicable existing legal norms and requirements can be achieved. Nothing in this document shall create legal rights nor impose legal obligations towards third parties. We however recall that it is the duty of any natural or legal person to comply with laws – whether applicable today or adopted in the future according to the development of AI. These Guidelines proceed on the assumption that **all legal rights and obligations that apply to the processes and activities involved in developing, deploying and using AI systems remain mandatory and must be duly observed**.

### Ethical AI

Achieving Trustworthy AI requires not only compliance with the law, which is but one of its three components. Laws

**Trustworthiness** is a prerequisite for people and societies to develop, deploy and use AI systems. Without AI systems – and the human beings behind them - **being demonstrably worthy of trust.**

**Trustworthy AI has three components, which should be met throughout the system's entire life cycle:**

- it should be **lawful**, complying with all applicable laws and regulations
- it should be **ethical**, ensuring adherence to ethical principles and values
- it should be **robust**, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm

# AI systems do not operate in a lawless environment

- **EU primary law** - the Treaties of the European Union and its Charter of Fundamental Rights
- **EU secondary law** - the General Data Protection Regulation/ the Product Liability Directive/ the Regulation on the Free Flow of Non-Personal Data/ anti-discrimination Directives/ consumer law and Safety and Health at Work Directives
- **UN Human Rights treaties and the Council of Europe conventions** - the European Convention on Human Rights
- **and numerous EU Member State laws**

## Chapter II key guidance:

- Systematically train a new generation of experts in AI ethics
- Facilitate the traceability and auditability of AI systems
- Involve stakeholders throughout the AI system's lifecycle. Foster training and education so that all the stakeholders are aware of and trained in Trustworthy AI

## Chapter III key guidance:

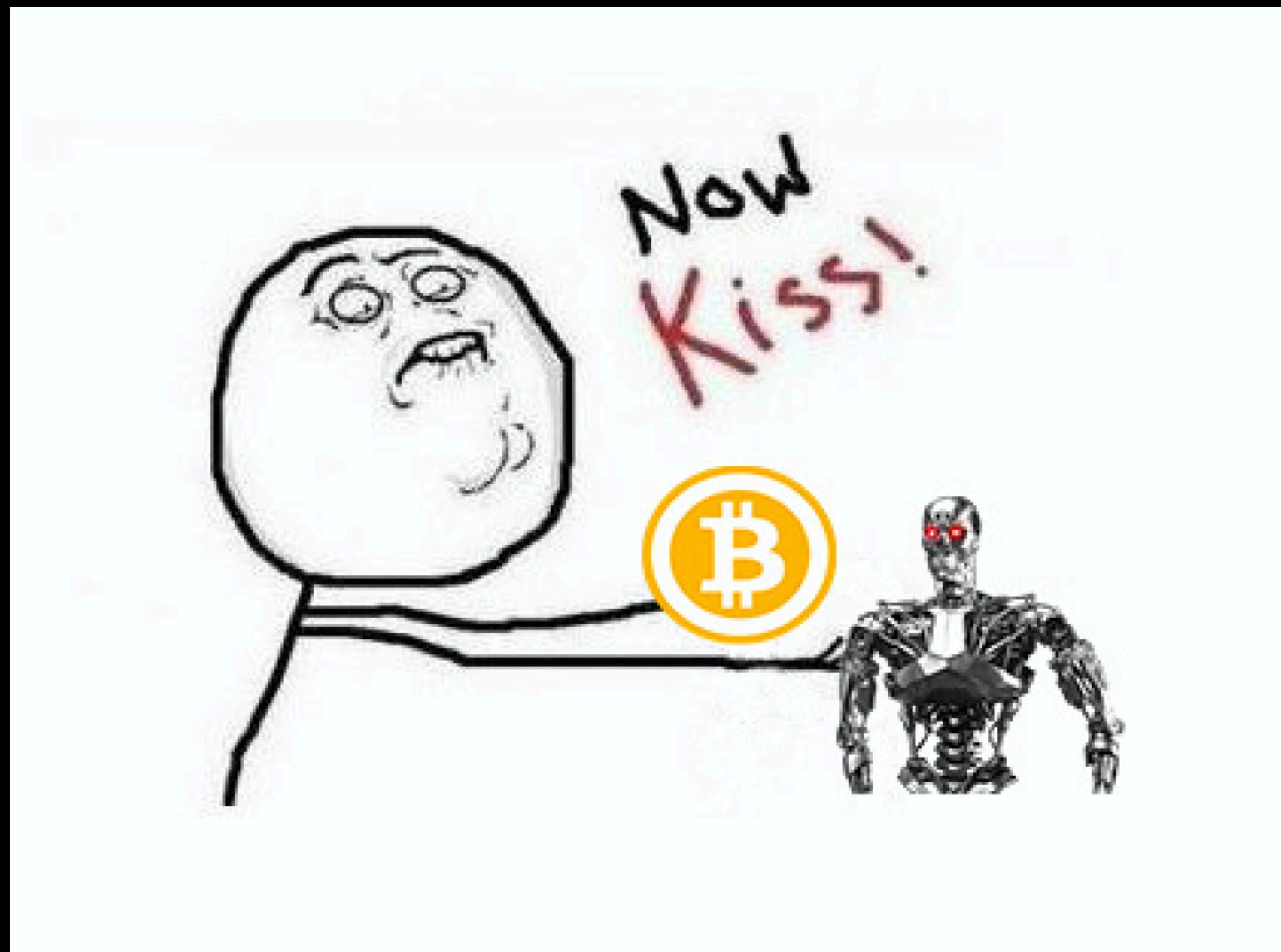
- Adopt a Trustworthy AI assessment list when developing, deploying or using AI systems, and adapt it to the specific use case in which the system is being applied
- Keep in mind that such an assessment list will never be exhaustive ^\\_(ツ)\_/^-

**«These Guidelines proceed on the assumption  
that all legal rights and obligations that  
apply to the processes and activities involved  
in developing, deploying and using AI systems  
remain mandatory and must be duly observed»**

**«These Guidelines proceed on the assumption  
that all legal rights and obligations that  
apply to the processes and activities involved  
in developing, deploying and using AI systems  
remain mandatory and must be duly observed»**



# How to live between two dystopias and still stay sane?



# Solution

---

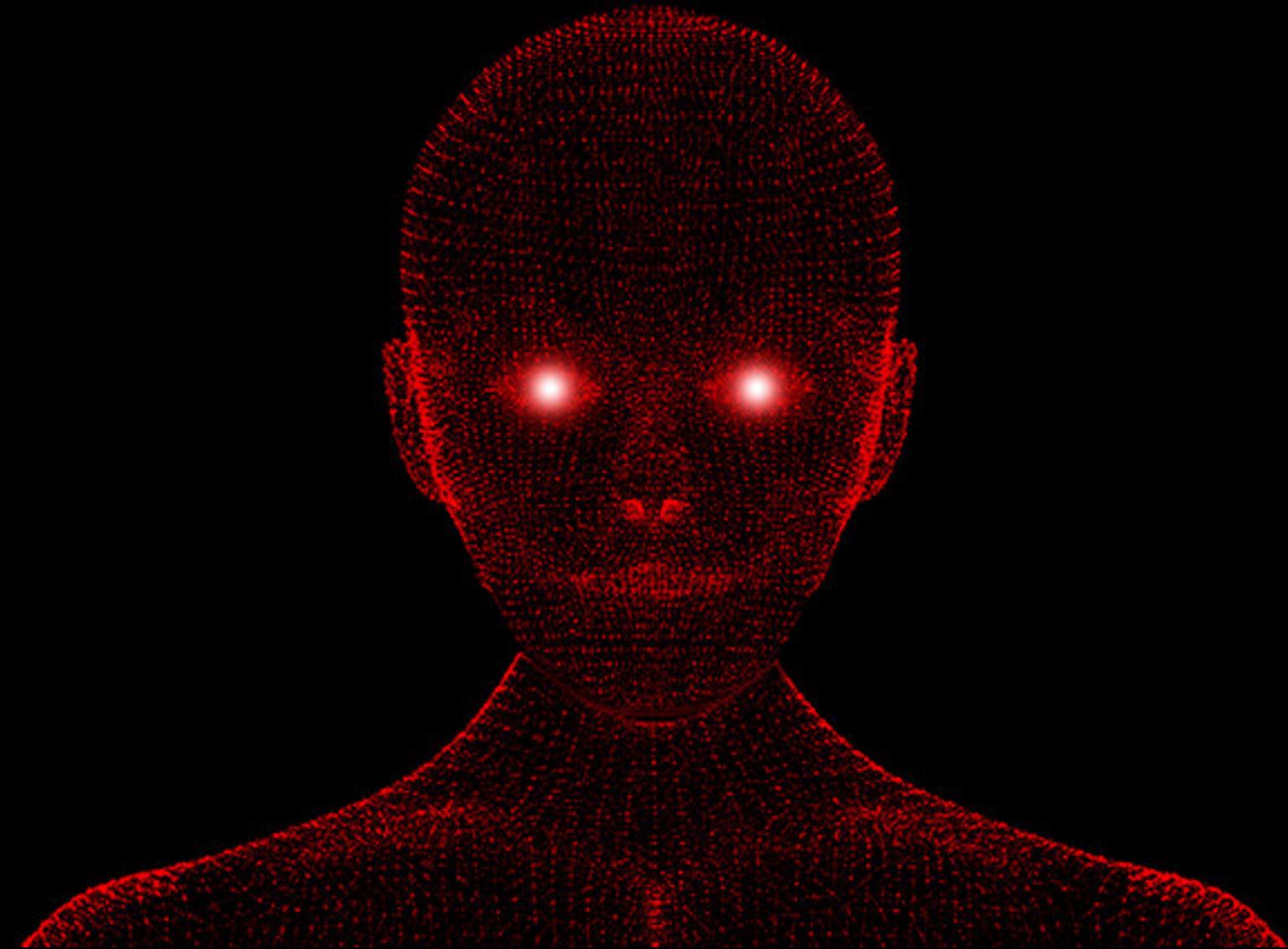
- Build new tools that would enable us to provide **censorship-resistant** AI tech (**Bitcoin L3**)
- Use cryptography for increasing privacy
- Don't fall for the scary Government/ ethical fairy tales
- Don't judge 'AI' before it's even there
- Follow the white rabbit



# Tools

- Complex CTLV/CSV contracts
- HTLCs (hash-timelocked contracts)
- Multi-party ECDSA and Shamir secret sharing (while we are waiting for Schnorr's signatures)
- SIGHASH\_NOINPUT-based contracts
- Zero-knowledge proofs
- Onion routing in LN
- Probabilistic checkable proofs
- Better homomorphic cryptography schemes
- TLV extensions in LN gossip protocol and P2P messages

# Thank you!



**Olga Ukolova, MD**  
[ukolova@pandoracore.com](mailto:ukolova@pandoracore.com)

- COO at Pandora Core
- Co-author of The #FreeAI Manifesto
- Specialist in Cognitive science