

WEBINAR

# Safeguarding open data: cybersecurity essentials and skills for data providers

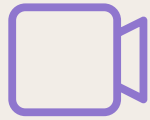
The logo for Data Europa Academy is located in the bottom left corner. It consists of the text "data." on the first line, "europa" on the second line, and "academy" on the third line. The text is white with small colored dots (yellow, blue, orange) above the letters 'a', 'e', and 'a' respectively. The logo is set against a dark blue circular background.

data.  
europa  
academy

18 October 2024

10:00 – 11:30 CEST

# Rules of the game



The webinar will be recorded and published on the data.europa academy



For questions, please use the ClickMeeting chat.



Please reserve 3 min after the webinar to help us improve by filling in our feedback form



# Introduction



**Inmaculada Farfan  
Velasco**  
Data.europa academy,  
Publications Office of the EU



**Roeland de Koning**  
Director Public Security,  
Capgemini



**Amanda Brincat**  
Data Security Consultant,  
Capgemini



**Prokopios Drogkaris**  
Deputy Data Protection Officer,  
European Union Agency for  
Cybersecurity



# Agenda

10.05 – 10.10

Opening and introduction – *Inmaculada Farfan Velasco*

10.10 – 10.40

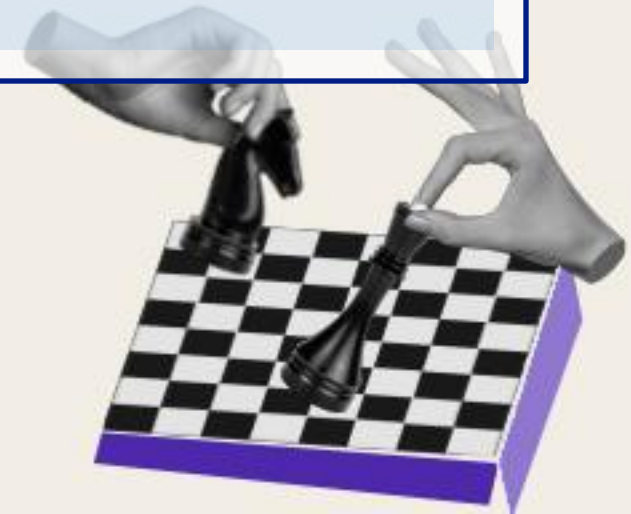
'Safeguarding Open Data' – *Roeland de Koning & Amanda Brincat*

10.40 – 11.10

'Striking a balance between open data and data protection' – *Prokopios drogkaris*

11.10 – 11.30

Q&A session and closing remarks



# SAFEGUARDING OPEN DATA

UNDERSTANDING THE RISKS,  
BALANCING OPENNESS  
AND CYBER SECURITY

Amanda Brincat and Roeland de Koning



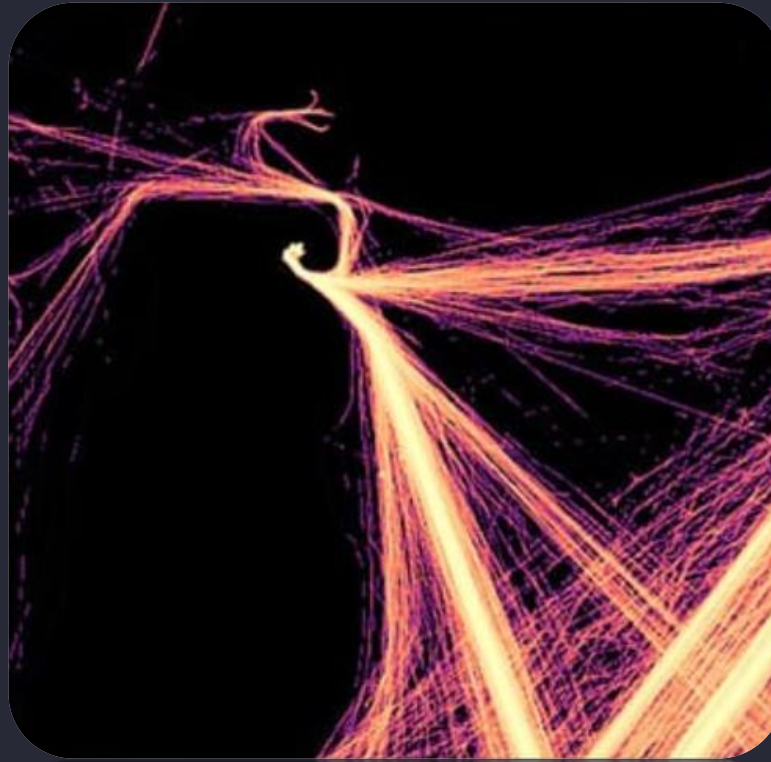


## AGENDA

- 01 The Openness - Security Paradox
- 02 Threats to Open Data
- 03 When Things go Wrong: Disinformation Case
- 04 Navigating the Data Publishing Dilemma
- 05 Responsible Data Sharing



# OPEN DATA IS BEAUTIFUL





# DATASETS AND ESPECIALLY HIGH-VALUE DATASETS ARE CRITICAL FOR NUMEROUS SERVICES AND NEEDS TO BE PROTECTED



## What is Open Data

Open data is data that anyone can access, use and share. Governments, businesses and individuals can use open data to bring about social, economic and environmental benefits

## Earth Observation & Environment

Satellite imagery and environmental data such as air quality and land use.

## Statistics

Aggregated data on demographics, economics, and social indicators.

## Geospatial

Data related to locations, maps, and geographical features.

## Meteorological

Weather and climate-related data, including forecasts and historical trends.



## Companies

Information on business entities, such as registration details and financials.

## Mobility

Transportation data, including traffic flows and public transit schedules.

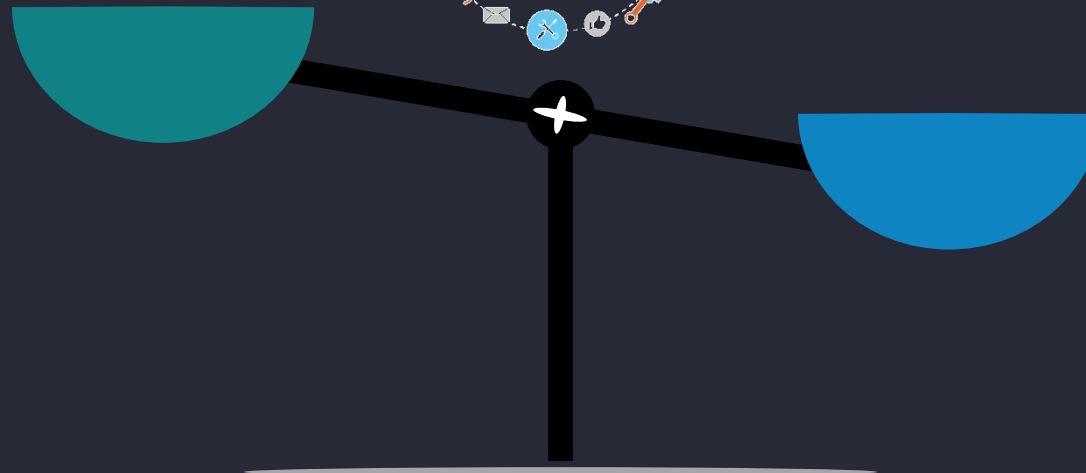




# OPENNESS AND CYBERSECURITY: CAN WE HAVE BOTH?

## Open data

- Gives access to public information
- Helps to make governments more transparent
- Provides the evidence that public money is being well spent and policies are being implemented
- Is opening up new opportunities for businesses to connect with customers.
- Helps us protect our planet by providing detailed early warning



## Cybersecurity

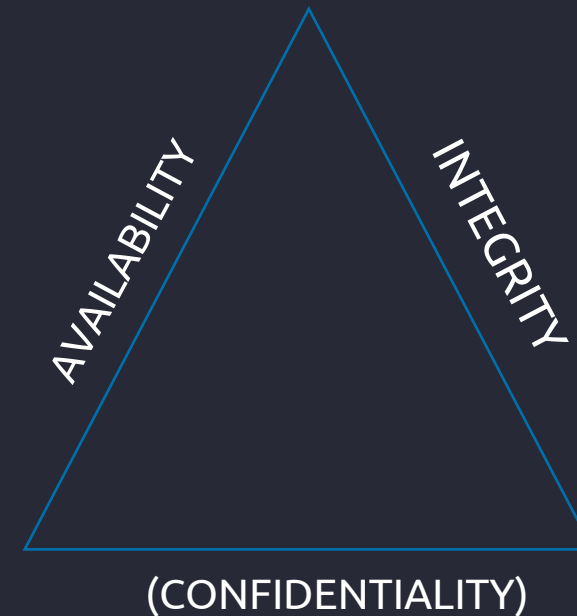
- Prevents vulnerable data to:
  - Tampering,
  - Unauthorized access,
  - Misuse.
- Protects the integrity and confidentiality of datasets
- Ensures trust, accuracy, and the continued usability

# IT SOUNDS LIKE A CONTRADICTION... HOW CAN SOMETHING BE OPEN AND SECURE?



The CIA cybersecurity triangle:

- Protect **confidentiality** by controlling access,
- Ensure **integrity** by preventing unauthorized changes,
- Maintain **availability** by keeping data reliable and accessible to the public.



**Open data requires a focus on integrity and availability**



# THREAT ACTORS CAN EXPLOIT VULNERABILITIES TO TARGET OPEN DATA BEFORE IT IS PUBLISHED



## Supply Chain Targeting

- Attackers increasingly target open-source projects by introducing malicious code into commonly used libraries or datasets.
- *These modifications can go unnoticed until they are integrated into broader systems, making them effective entry points for compromising data integrity.*

## Information Manipulation

- Introducing false values or reshaping narratives by exploiting the transparency of open data initiatives.  
*Manipulating open data before it is consumed by end-users can erode public trust.*

## Data Poisoning

- This tactic involves injecting misleading or incorrect data into datasets, particularly those used for AI training.
- *By contaminating the data, attackers can alter AI model behaviors, skewing outcomes and potentially causing business or operational failures.*

# AFTER VULNERABILITIES HAVE BEEN EXPLOITED THREAT ACTORS CAN MISUSE EXPOSED DATA THROUGH FOUR MAIN TACTICS



These threats are referred to as Post-Exploitation within the lifecycle of a cyberattack.



## Data Tampering and Integrity Attacks

Altering published datasets to mislead decision-makers.



## Data Re-identification

Reconstructing identities from anonymized data.



## AI Manipulation

Introducing subtle changes to input data used in AI systems leading to misclassifications.



## Sensitive Data Leakage

Intentional exposure of personal or sensitive information.



(THREAT REPORT 2024)



# (CYBER) DISINFORMATION OPERATES AS BOTH A PRE-EXPLOITATION AND POST-EXPLOITATION TACTIC

“Disinformation is **false or misleading** content that is spread with an **intention** to deceive or secure economic or political gain, and which may cause **public harm**.”



NPR  
<https://www.npr.org> › 2022/03/28 › renewable-energy-...

## Misinformation is stopping renewable energy projects

Mar 28, 2022 — The spread of **misinformation** about solar and wind **energy** is leading some states and counties to restrict or even reject projects.



EU DisinfoLab  
<https://www.disinfo.eu> › publications › the-war-climate-c...

## The war, climate change, the energy crisis, and always ...

Jan 30, 2023 — **2022** was marked by **disinformation** that preyed on the sense of financial stability and security – from the **energy** crisis to xenophobic fears.



European Commission  
<https://commission.europa.eu> › news › 2022-state-energ...

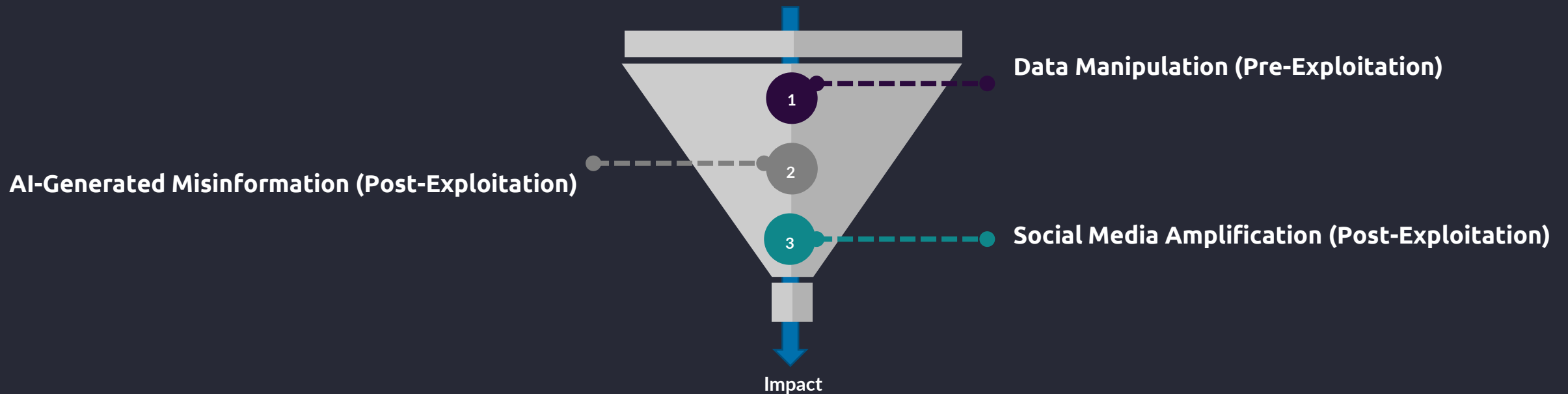
## 2022 State of the Energy Union report highlights challenges

Oct 18, 2022 — 2022 has seen **turbulence in energy markets, price volatility and energy insecurity** across the world, which has had a huge impact on the EU's ...



# THE 2022 ENERGY DATA MANIPULATION HIGHLIGHTS HOW DISINFORMATION CAN EXACERBATE AN EXISTING ENERGY CRISIS

## Energy Grid Disinformation (2022)



### Supply Chain Disruptions

Panic-buying led to temporary shortages of fuel and generators.

### Public Anxiety

Citizens were left confused and fearful about the energy grid's stability.

### Economic Consequences

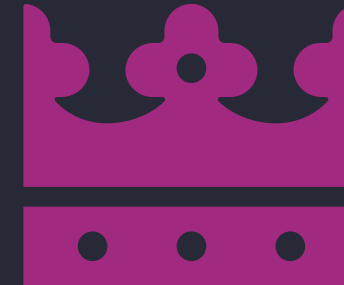
Induced fluctuations in the stock market and increased energy prices.



# NAVIGATING THE DATA PUBLISHING DILEMMA: WHAT SHOULD AND SHOULDN'T BE SHARED?

## Prevent unwanted effects of publishing open data

- **Protect Privacy**
- **Prevent Market disruption**
- **Consider Data Combination Risks**
- **Compromise the safety or security**
- **Cause commercial or economic harm**
- **Compromising rights of others**
- **Jeopardise national security**



## Identify "Crown Jewel" Data

- **Critical and Sensitive Information**
- **Personal and Sensitive Public Data**
- **National Security Data**

Source : [WWW.npsa.gov.uk/](http://WWW.npsa.gov.uk/) CPNI



# RESPONSIBLE DATA SHARING NEEDS TO ASSESS OPENNESS AND SECURITY OF THE DATA

## 6 Elements for assessing Openness and Security

Identify sensitive data

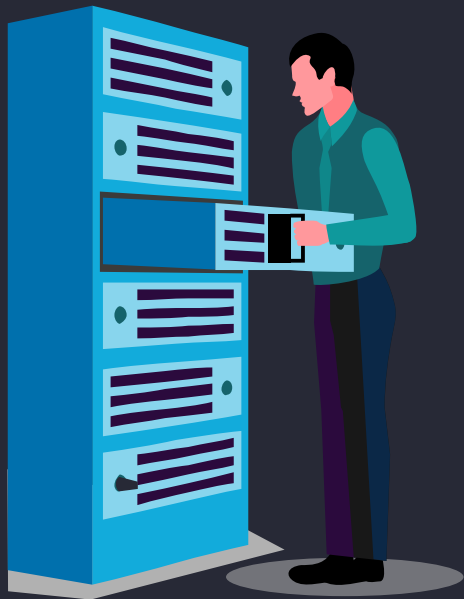
Risk Assessment

Data sharing agreement

Security management

Risk mitigation

Data governance



## Some Tips for risk mitigation

### Implement Effective Risk Mitigation Strategies

- Remove sensitive data subsets to minimize exposure.
- Provide summary information instead of detailed datasets to limit granularity.

### Leverage Privacy Enhancing Technologies

- Utilize PET technologies like data anonymization and pseudonymization to safeguard individual identities in shared datasets.

### Monitor Access and Usage

- If possible, apply user registration to control access and track and manage who can view sensitive data.

**We need to encourage responsible sharing of data building on a current “culture of sharing”**

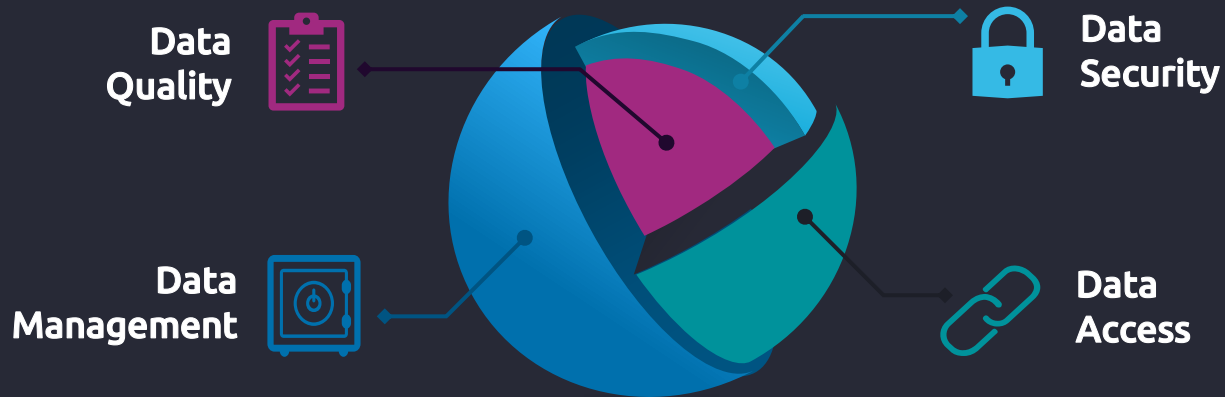




# RESPONSIBLE DATA SHARING IS ENABLED BY DATA GOVERNANCE

Data Governance builds trust in data, a crucial asset in today's data-enabled decision-making. Trustworthy data enhances decision accuracy, reduces risks, and fosters compliance with data protection regulations.

## The 4 Pillars to Data Governance



## From Open Data to Data Spaces

### Data Governance Act

- Overarching horizontal governance **framework for EU data spaces**
- There is a need for data spaces to share data in a **controlled manner**
- Tools for **data pooling, access, sharing, and managing access rights.**

Reference : Open data Institute (framework of essential data practices for trustworthiness of data)

**Openly and responsibly data sharing can be used to help society make better decisions**



Thank you for your attention!

Please save your questions for the Q&A



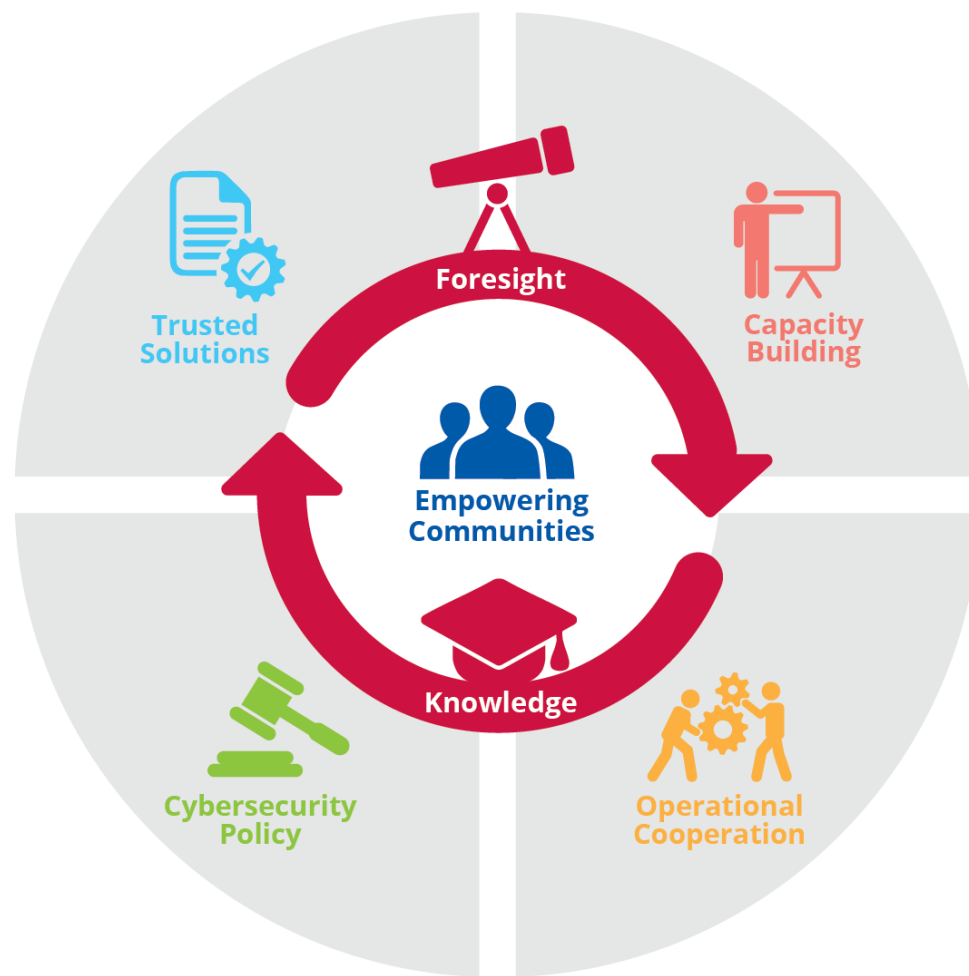
EUROPEAN UNION AGENCY  
FOR CYBERSECURITY

# STRIKING A BALANCE BETWEEN OPEN DATA AND DATA PROTECTION

Dr. Prokopios Drogkaris – ENISA  
Cybersecurity Expert – Deputy DPO

18 | 10 | 2024

# ENISA STRATEGIC OBJECTIVES





# WHAT IS PSEUDONYMISATION?

## GDPR (Article 4(5)):

- the processing of personal data in such a manner that the personal data **can no longer be attributed to a specific data subject without the use of additional information**, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person



# WHAT IS ANONYMISATION?

## GDPR (recital 26)

- anonymous information refers to information which **does not relate to an identified or identifiable natural person**

## ISO/TS 25237:2017

- process by which personal data is **irreversibly altered** in such a way that a data subject can **no longer be identified directly or indirectly**, either by the data controller alone or in collaboration with any other party



# ANONYMISATION ≠ PSEUDONYMISATION

**Pseudonymous data = Personal Data**

**Anonymous data (impossible to identify individuals) ≠ Personal Data**

**Common Misperception:**

*“If I remove direct identifiers from the dataset (e.g. name, address, social security number, etc.), the data is anonymised”*

# WHAT CAN GO WRONG?

## Researchers reverse Netflix anonymization

Robert Lemos, SecurityFocus 2007-12-04

In a dramatic demonstration of the privacy dangers of databases that collect consumer habits, two researchers from the University of Texas at Austin have shown that a handful of movie ratings can identify a person as easily as a Social Security number.

The researchers -- graduate student Arvind Narayanan and professor Vitaly Shmatikov, both from the Department of Computer Sciences at the University of Texas at Austin -- claim to have identified two people out of the nearly half million anonymized users whose movie ratings were released by online rental company Netflix last year. The company published the large database as part of its \$1 million Netflix Prize, a challenge to the world's researchers to improve the rental firm's movie-recommendation engine.

"Releasing the data and just removing the names does nothing for privacy," Shmatikov told SecurityFocus. "If you know their name and a few records, then you can identify that person in the other (private) database."

**"Releasing the data and just removing the names does nothing for privacy. If you know their name and a few records, then you can identify that person in the other (private) database."**

Vitaly Shmatikov, Professor of Computer Science,  
University of Texas at Austin

Regulation (EU) 2018/1807 on a framework for the free flow of non-personal data in the European Union\*

- mentions "*...market developments and technological developments which might expand the possibilities for deanonymising data...*"

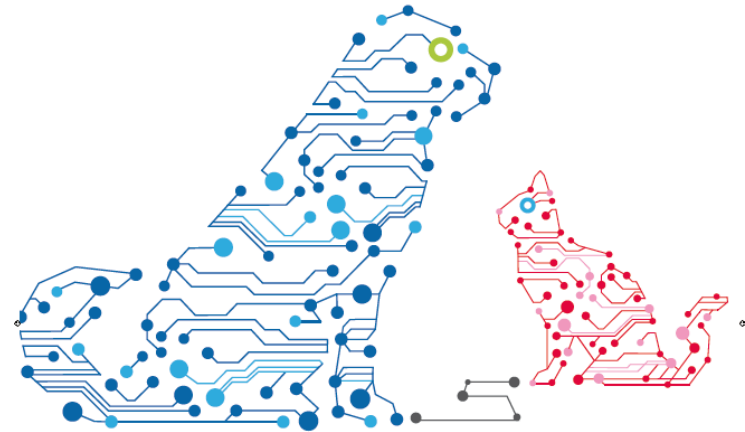
\* <http://data.europa.eu/eli/reg/2018/1807/oj>



# PRIVACY ENHANCING TECHNOLOGIES

- Anonymisation & pseudonymisation techniques
- Data masking and privacy preserving computations
  - (e.g. homomorphic encryption, secure multiparty computations, synthetic data,..)
- Access, communication & storage
  - (e.g. end-to-end encryption, proxy/onion routing,..)
- Privacy-enhancing access control, authorisation and authentication (
  - e.g. privacy-enhancing attribute-based credentials, zero knowledge proof, ..)
- Transparency, intervenability & user control tools
  - (privacy policies, privacy icons, sticky policies, privacy dashboards, ..)
- Consent management tools

Privacy Enhancing Technologies  
protect your online privacy

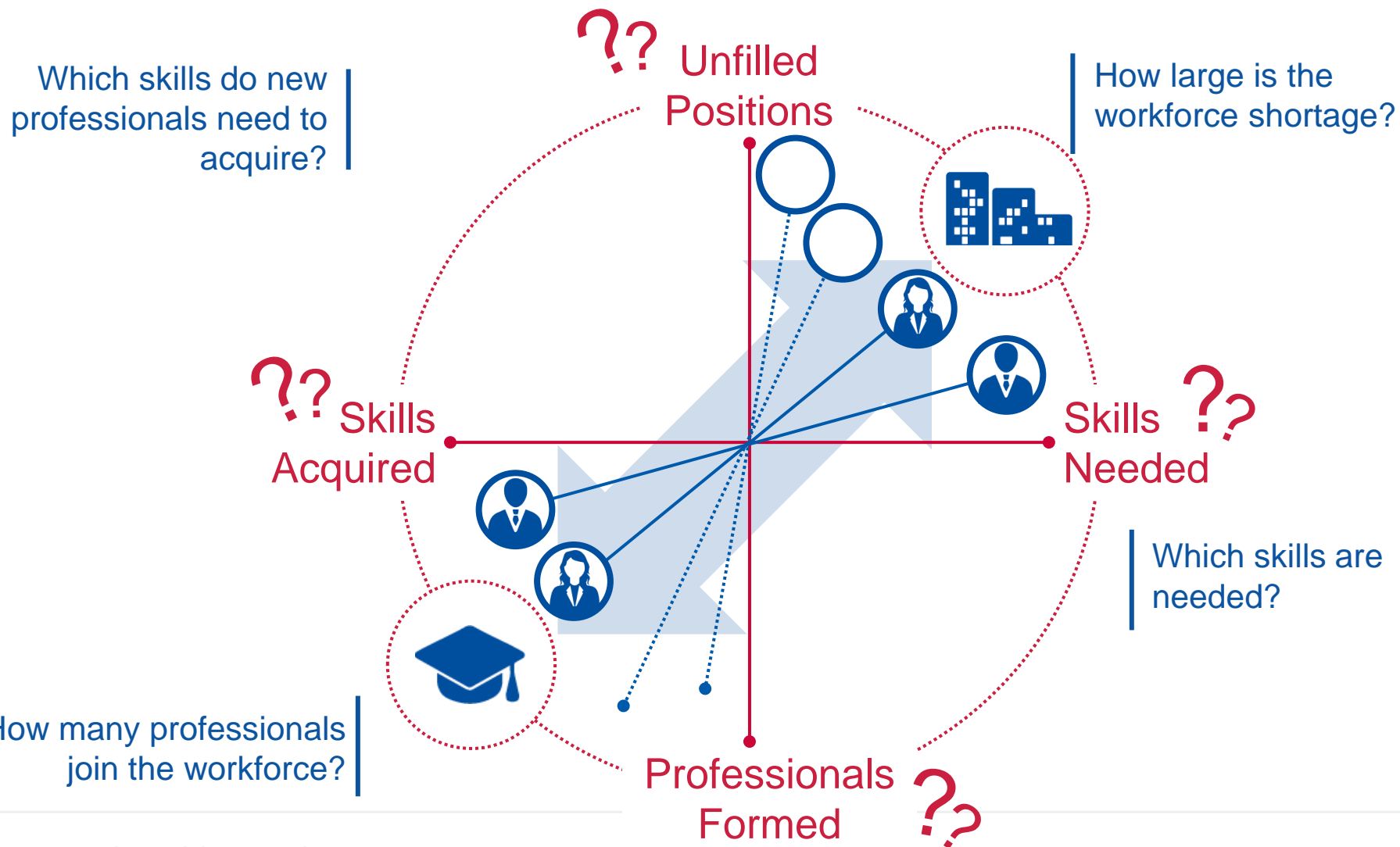


## Time to adopt PETs!

Privacy Enhancing Technologies (PETs) help to protect online privacy following the simple approach "reduce, protect, detect".  
The future starts now: make it a habit, adopt PETs.

Source: ENISA, Data protection engineering (2022)

# CYBERSECURITY SKILLS GAP AND SHORTAGE



# ECSF CYBERSECURITY PROFILES



**Chief Information Security Officer (CISO)**



**Cyber Incident Responder**



**Cyber Legal, Policy and Compliance Officer**



**Cyber Threat Intelligence Specialist**



**Cybersecurity Architect**



**Cybersecurity Auditor**



**Cybersecurity Educator**



**Cybersecurity Implementer**



**Cybersecurity Researcher**



**Cybersecurity Risk Manager**



**Digital Forensics Investigator**



**Penetration Tester**

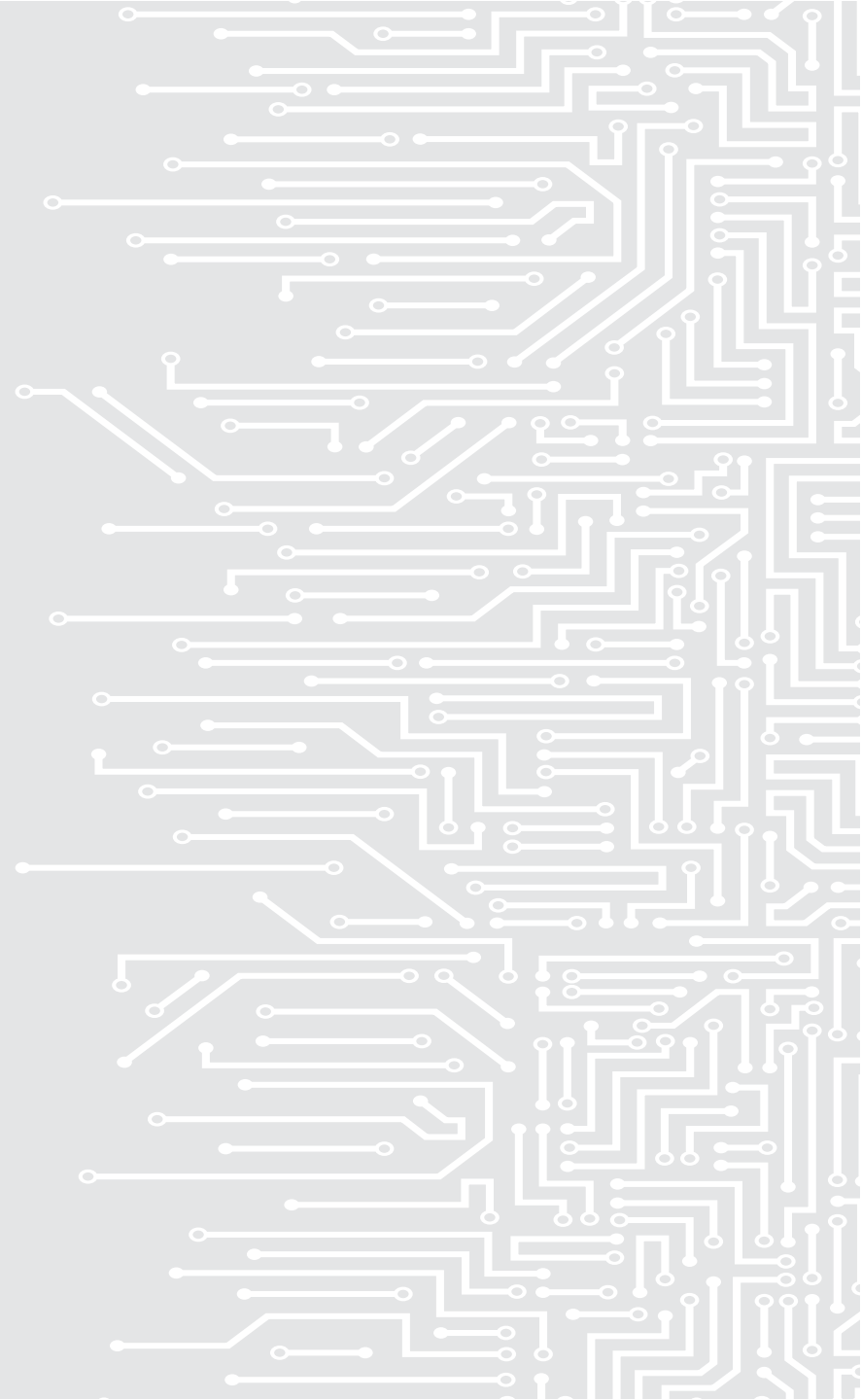
# THANK YOU FOR YOUR ATTENTION

Agamemnonos 14, Chalandri 15231  
Attiki, Greece

 +30 28 14 40 9711

 [info@enisa.europa.eu](mailto:info@enisa.europa.eu)

 [www.enisa.europa.eu](http://www.enisa.europa.eu)



# Q&A



**Inmaculada Farfan  
Velasco**  
Data.europa academy,  
Publications Office of the EU



**Roeland de Koning**  
Director Public Security,  
Capgemini



**Amanda Brincat**  
Data Security Consultant,  
Capgemini



**Prokopios Drogkaris**  
Deputy Data Protection Officer,  
European Union Agency for  
Cybersecurity



Stay up-to-date on our  
2024 activities!

The logo for Data Europa Academy is located in the bottom left corner. It features the text "data." on the first line, "europa" on the second line, and "academy" on the third line. The number "30" is positioned between "europa" and "academy". The text is white with small colored dots (yellow, blue, orange) above the letters. The logo is set against a dark blue circular background, which is itself inside a larger purple circular shape.

data.  
europa  
30  
academy

WORKSHOP

# How to use open data for your research

**data.**  
**europa**  
academy

25 October 2024  
10.00 – 11.30 CEST



**data.**  
**europa**  
academy



# Your opinion is important to us





# Thank you!

