Privacy

Violet Shi
2020.11.10

General Data Protection Regulation

GDPR

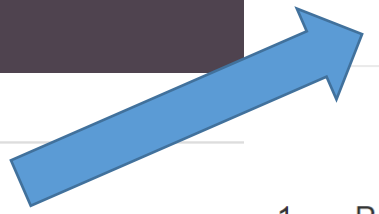
Welcome to gdpr-info.eu. Here you can find the official PDF of the Regulation (EU) 2016/679 (General Data Protection Regulation) in the current version of the OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018 as a neatly arranged website. All Articles of the GDPR are linked with suitable recitals. The European Data Protection Regulation is applicable as of May 25th, 2018 in all member states to harmonize data privacy laws across Europe. If you find the page useful, feel free to support us by sharing the project.

The **General Data Protection Regulation** (**GDPR**) is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA)

https://gdpr-info.eu/

Principles in processing personal data

Chapter 2	–	**Principles**
Article 5	–	Principles relating to processing of personal data
Article 6	–	Lawfulness of processing
Article 7	–	Conditions for consent
Article 8	–	Conditions applicable to child's consent in relation to information so services
Article 9	–	Processing of special categories of personal data
Article 10	–	Processing of personal data relating to criminal convictions and offe
Article 11	–	Processing which does not require identification



Art. 5 GDPR

Principles relating to processing of personal data

1. Personal data shall be:

(a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');

(b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');

(c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation');

(d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy');

https://gdpr-info.eu/

Data should be …

1. collected for specific, expressly stated and justified purposes and not treated in a new way that is incompatible with these purposes
   (principle of purpose limitation)

2. adequate, relevant and limited to what is necessary for fulfilling the purposes for which it is being processed
   (principle of data minimization)

3. processed in a lawful, fair and transparent manner
   (principle of legality, fairness and transparency)

Some definitions

Personal Data

(GDPR Article 4 (1))

Directly or indirectly linked to a person

Processing

(GDPR Article 4 (2))

Operations performed on personal data

Data controller

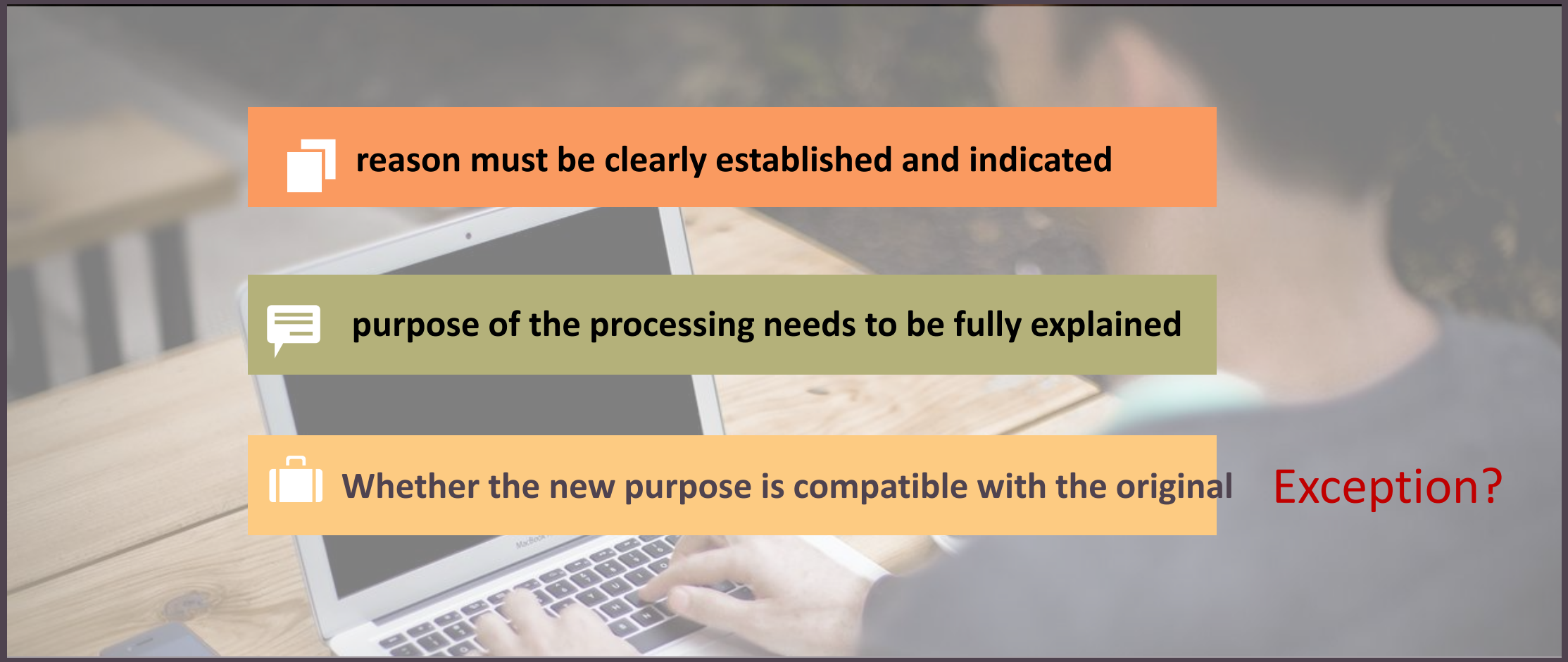
(GDPR Article 4 (7))

Bodies determines the purpose and means of processing

Purpose limitation

Data should be collected for specific, expressly stated and justified purposes, and not treated in a new way that is incompatible with these purposes

What to do?



- reason must be clearly established and indicated
- purpose of the processing needs to be fully explained
- Whether the new purpose is compatible with the original

Exception?

Problem: how to identify the exceptions?

scientific research: further process is considered to be compatible with the original purpose

Definition/Scope of scientific research?

1. Identify the risk of tax fraud based on person's activities in bank

2. Image recognition software diagnosing cancer in tumors

scientific research

Discover new knowledge or know-how.

Software life cycle:

development, application

scientific research

Differentiate between the development and the application of AI?

1. When the completed model is static (offline)

1. When the models develop and improve continuously as they are fed more personal data

Data minimization

Data should be adequate, relevant and limited to what is necessary for fulfilling the purposes for which it is being processed

What to do?

Controller

cannot use more personal data than is necessary

limits the intervention in a data subject's privacy

Benefit

✓ Minimize the risk of misleading models

Blackbox

Hard to predict what the algorithm will learn

Hard to define which data is necessary

Transparency in black box

processed in a lawful, fair and transparent manner

What to do?

Data subjects must be informed about how the information will be used.

the information must be easily available

1. impossible to explain how information is correlated

2. may reveal commercial secrets

Tools for good privacy

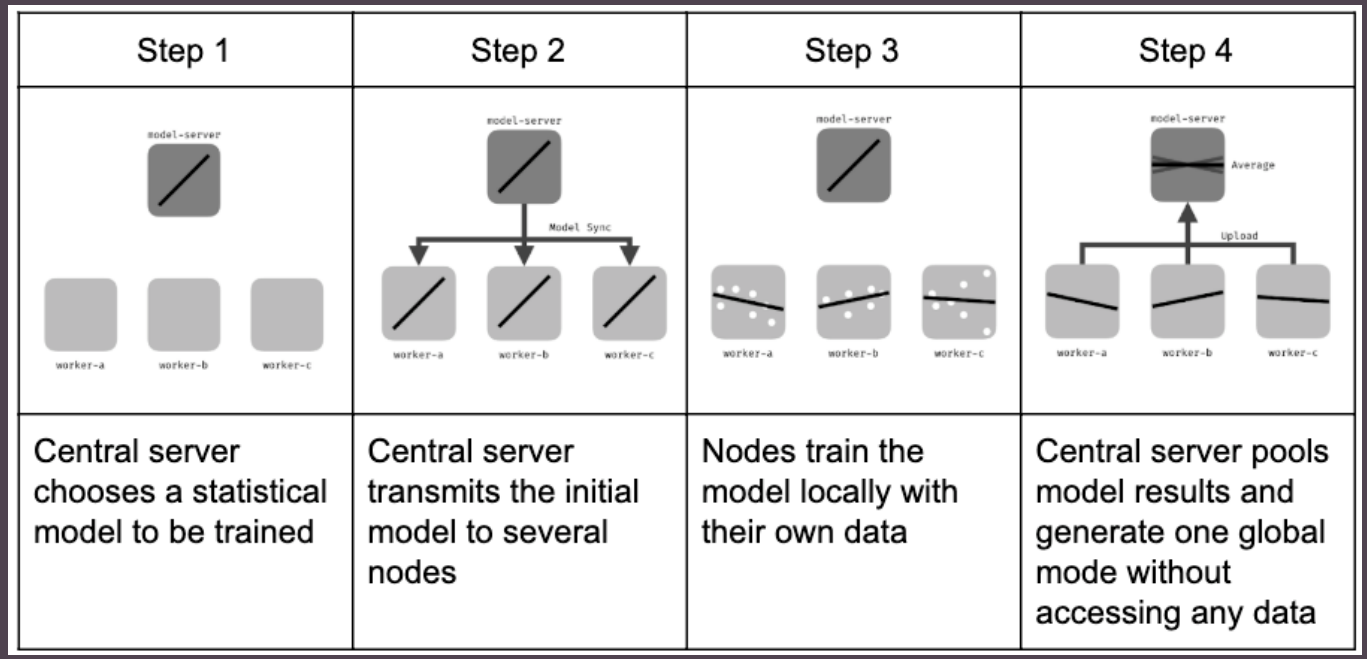
Reducing the need for training data

Generative Adversarial Network (GAN):

generating huge volumes of high quality, synthetic training data

Federated Learning:

improve an existing model, on a large number of users, without having to share the users' data



https://en.wikipedia.org/wiki/Generative_adversarial_network

https://en.wikipedia.org/wiki/Federated_learning

Protect privacy without reducing data basis

Differential privacy

contain deliberately-generated “noise”

Homomorphic encryption:

enables the processing of data whilst it is still encrypted

Transfer learning:

utilize existing models that solve similar tasks

https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf,
https://arxiv.org/abs/1412.7584

Reference

1. Artificial intelligence and privacy Report, January 2018
2. General Data Protection Regulation
3. The Algorithmic Foundations of Differential Privacy
4. Differential Privacy and Machine Learning: a Survey and Review

THANK YOU