


# T-Identity Protector.

Description of the conception and possible use-case scenarios



## T-Identity Protector.

Problem area/range of features.



### Problem

- Legal restrictions that make it difficult for organizations to generate analyses without delay when needed.

### Solution

- Anonymization, pseudonymization and re-personalization regulated technically and procedurally based on requirements.



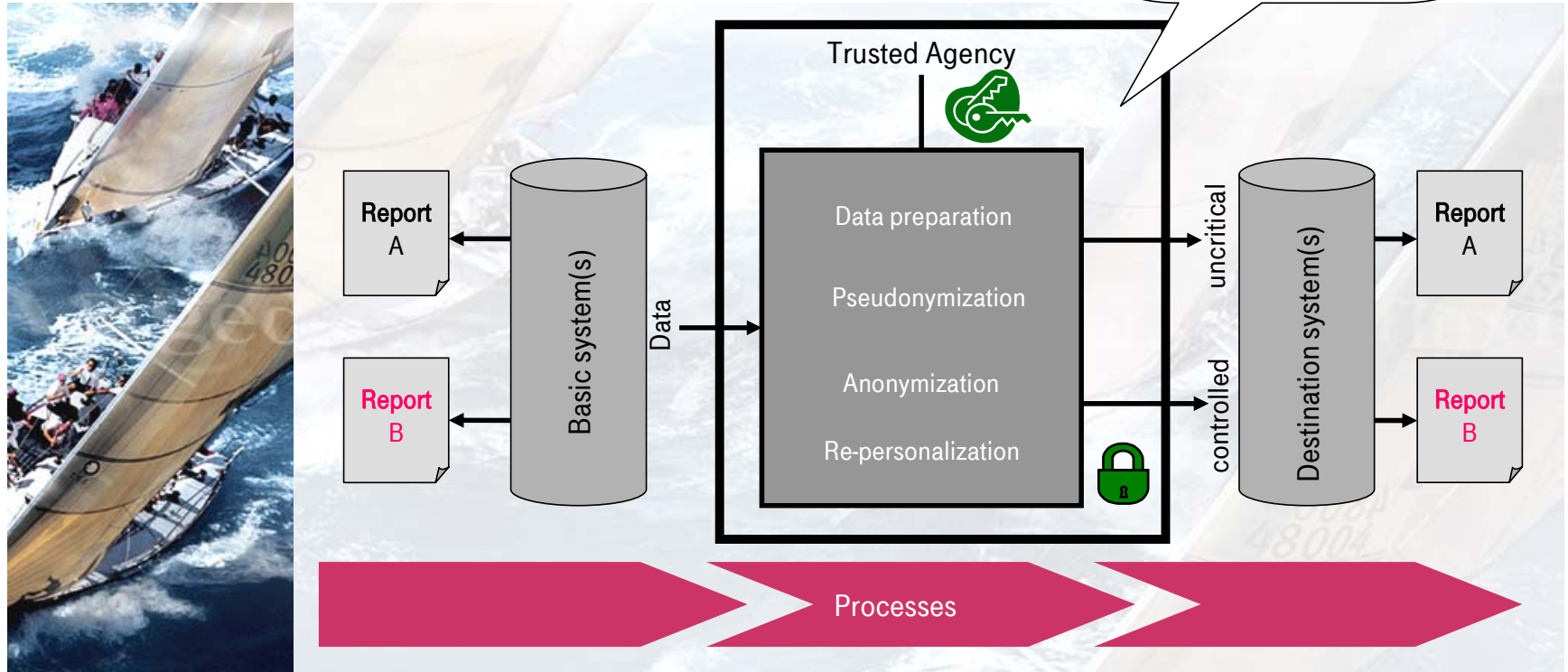
## T-Identity Protector.

Functionality, characteristic.



- Monitoring of data preparation of the basic systems
- Effective process for pseudonymization
- Conversion of identities into pseudonyms
- Conversion of pseudonyms into identities
- Controlled disclosure and release of identity
- Independent system under the control of a Trusted Agency
- Universal system between source and destination systems

## T-Identity Protector. Black box principle.



## T-Identity Protector. Pseudonymization.



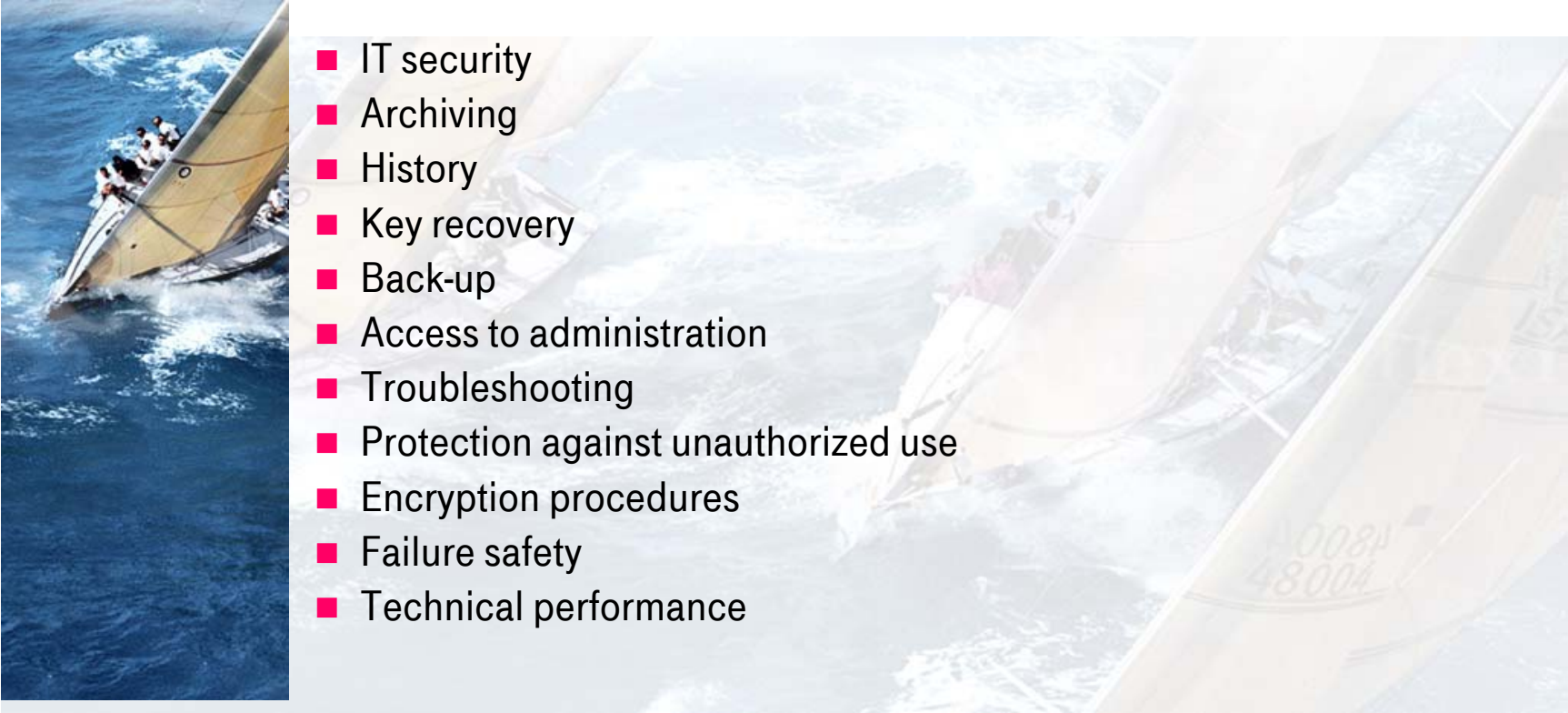
Pseudonymization is the changing of personal data using a matching rule in such a way that the individual entries regarding personal or factual conditions can no longer be assigned to an individual without knowledge or use of the matching rule.

To do so, identification datas, for example, are converted to a randomly selected code (the pseudonyms) via a mapping rule.

The aim of such a process is to be able to re-establish the personal reference only if needed and in compliance with previously defined general conditions.

Re-identification can sometimes also be left exclusively to the data subject.

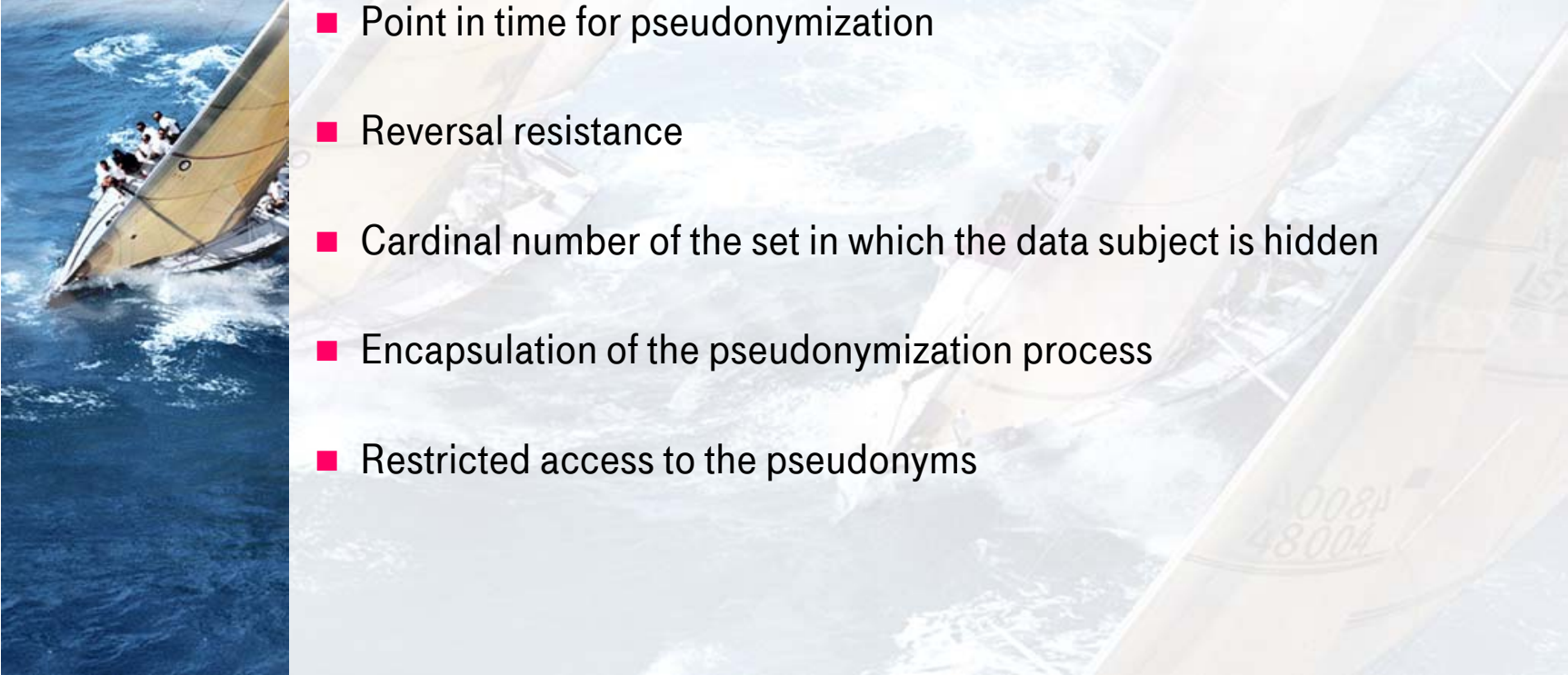
## T-Identity Protector. Technical requirements for the black box

- 
- IT security
  - Archiving
  - History
  - Key recovery
  - Back-up
  - Access to administration
  - Troubleshooting
  - Protection against unauthorized use
  - Encryption procedures
  - Failure safety
  - Technical performance



## T-Identity Protector.

### Pseudonymization – quality factors.

- 
- Point in time for pseudonymization
  - Reversal resistance
  - Cardinal number of the set in which the data subject is hidden
  - Encapsulation of the pseudonymization process
  - Restricted access to the pseudonyms

## T-Identity Protector. Risks + risk avoidance

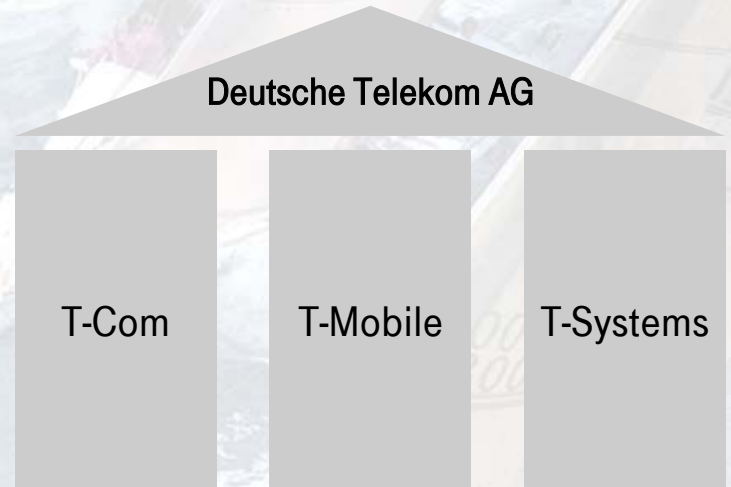
- All interfaces must be sufficiently secured
- The encryption level must also be secure and appropriate
- The human factor is always present



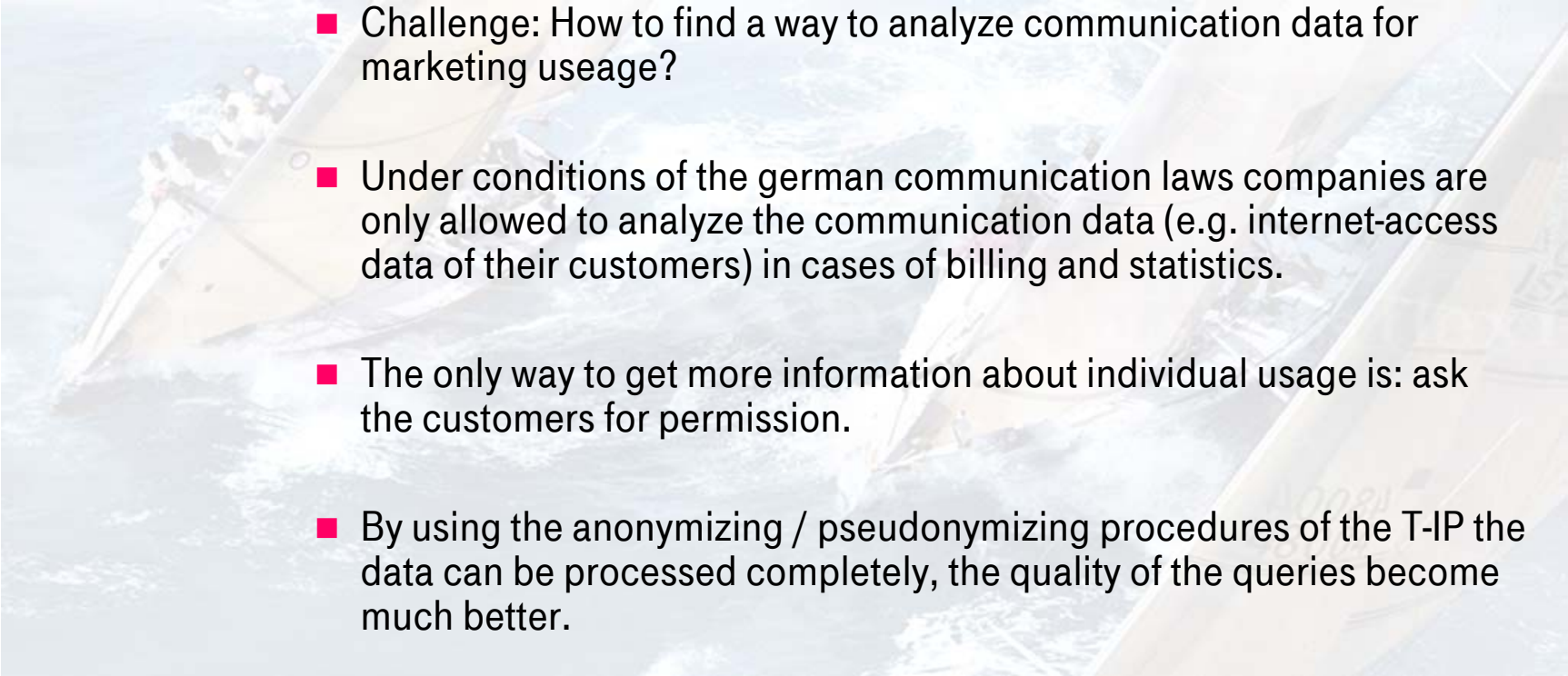


## T-Identity Protector. Use-case: skill management

- How to find special skills within the whole group ?
- Under restrictions of the german privacy law individual consents or agreements with the different works councils are required.
- By using the T-IP all data can be processed in a central area. The results of queries can be transmitted to the legal entities. Re-identification can be done in the entities.

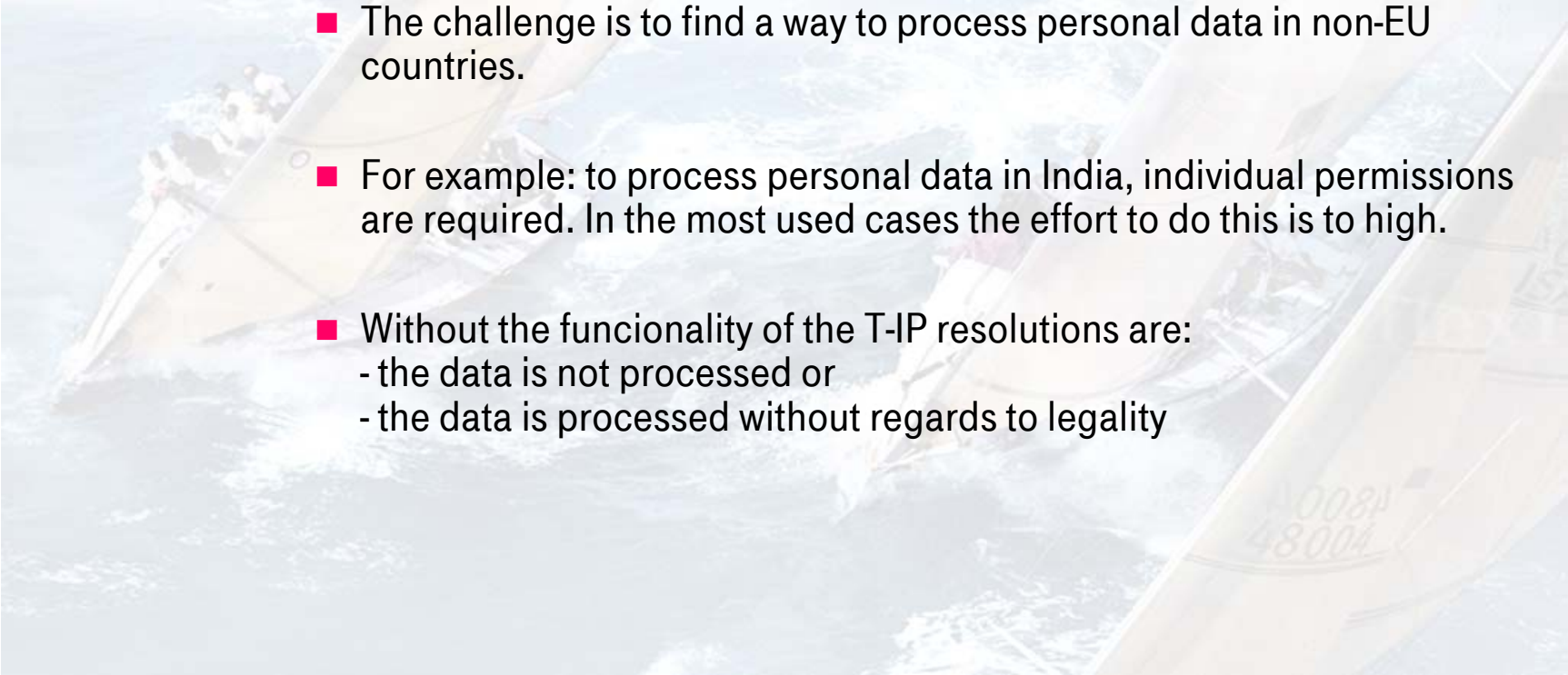


## T-Identity Protector. Use-case: Billing processes & marketing

- 
- Challenge: How to find a way to analyze communication data for marketing useage?
  - Under conditions of the german communication laws companies are only allowed to analyze the communication data (e.g. internet-access data of their customers) in cases of billing and statistics.
  - The only way to get more information about individual usage is: ask the customers for permission.
  - By using the anonymizing / pseudonymizing procedures of the T-IP the data can be processed completely, the quality of the queries become much better.

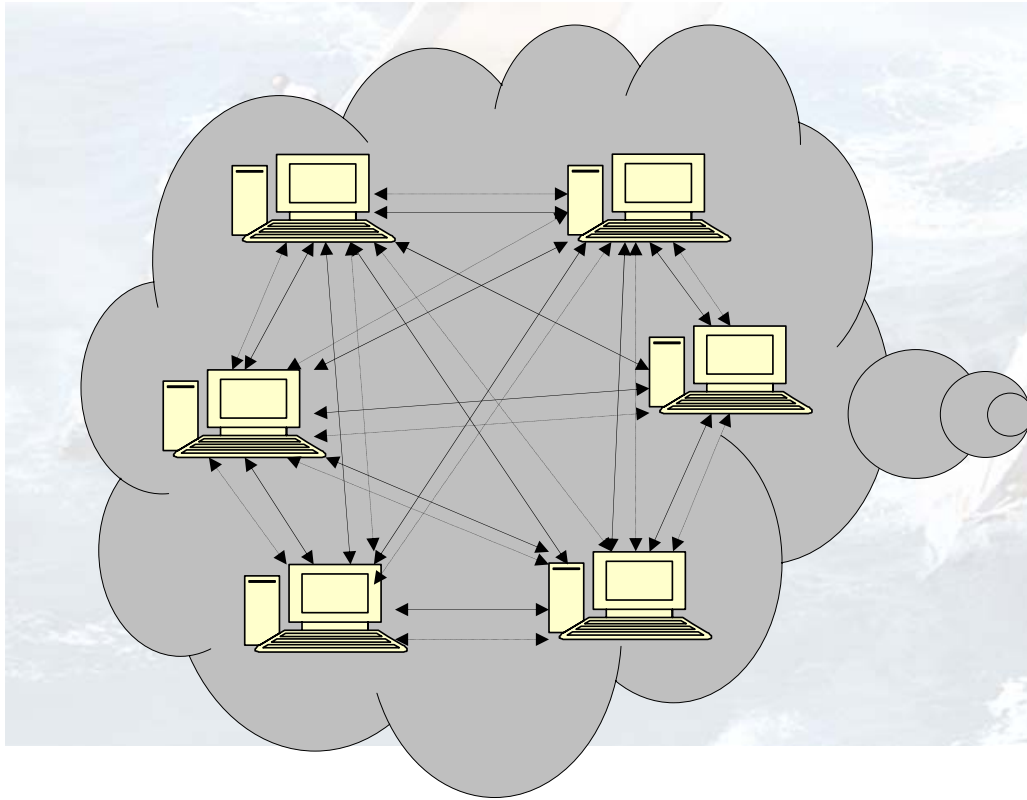
## T-Identity Protector.

### Use-case: Processing personal data in non-EU countries

- 
- The challenge is to find a way to process personal data in non-EU countries.
  - For example: to process personal data in India, individual permissions are required. In the most used cases the effort to do this is too high.
  - Without the functionality of the T-IP resolutions are:
    - the data is not processed or
    - the data is processed without regards to legality

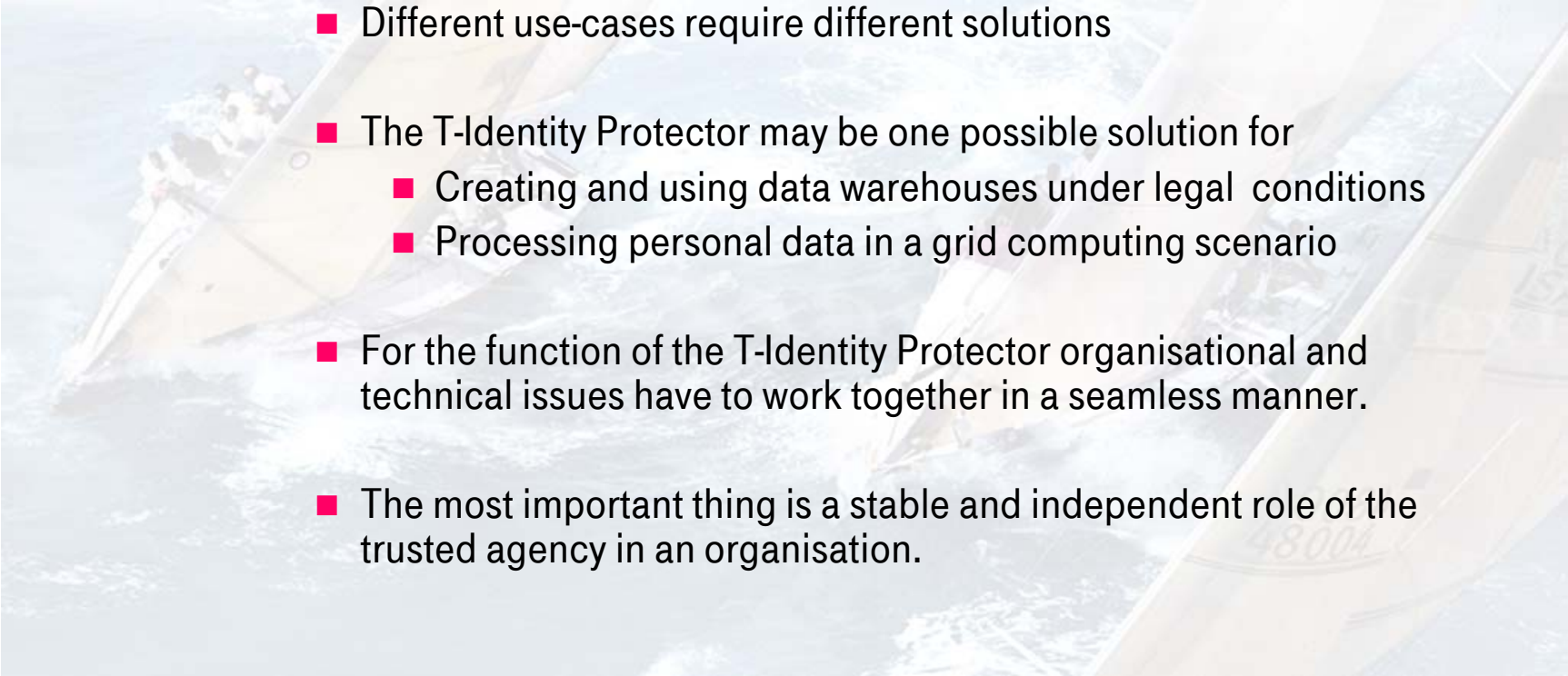


## T-Identity Protector. Use-case: Grid computing



- There is a contradiction between the restrictions of the (german) privacy laws and the philosophy of grid computing.
- Without individual permissions the transfer of personal data to non-EU countries is not allowed
- How to use the grid for processing personal data?

## Conclusion

- 
- Different use-cases require different solutions
  - The T-Identity Protector may be one possible solution for
    - Creating and using data warehouses under legal conditions
    - Processing personal data in a grid computing scenario
  - For the function of the T-Identity Protector organisational and technical issues have to work together in a seamless manner.
  - The most important thing is a stable and independent role of the trusted agency in an organisation.

An aerial photograph of three sailboats racing on a blue sea with white foam. The sailboats have large, light-colored sails. The text "Thank You." is overlaid in red at the top center.

Thank You.

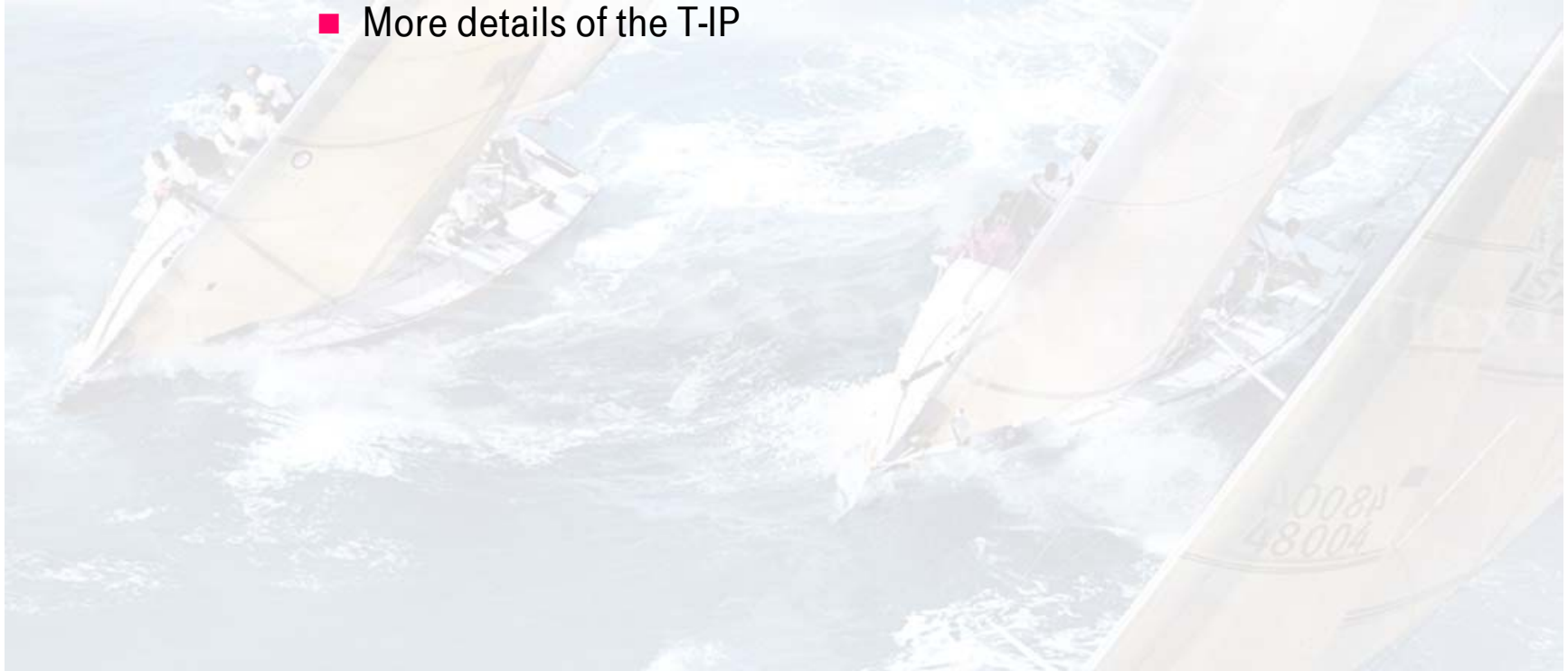
WagnerF@t-systems.com

.....T.....Systems.....



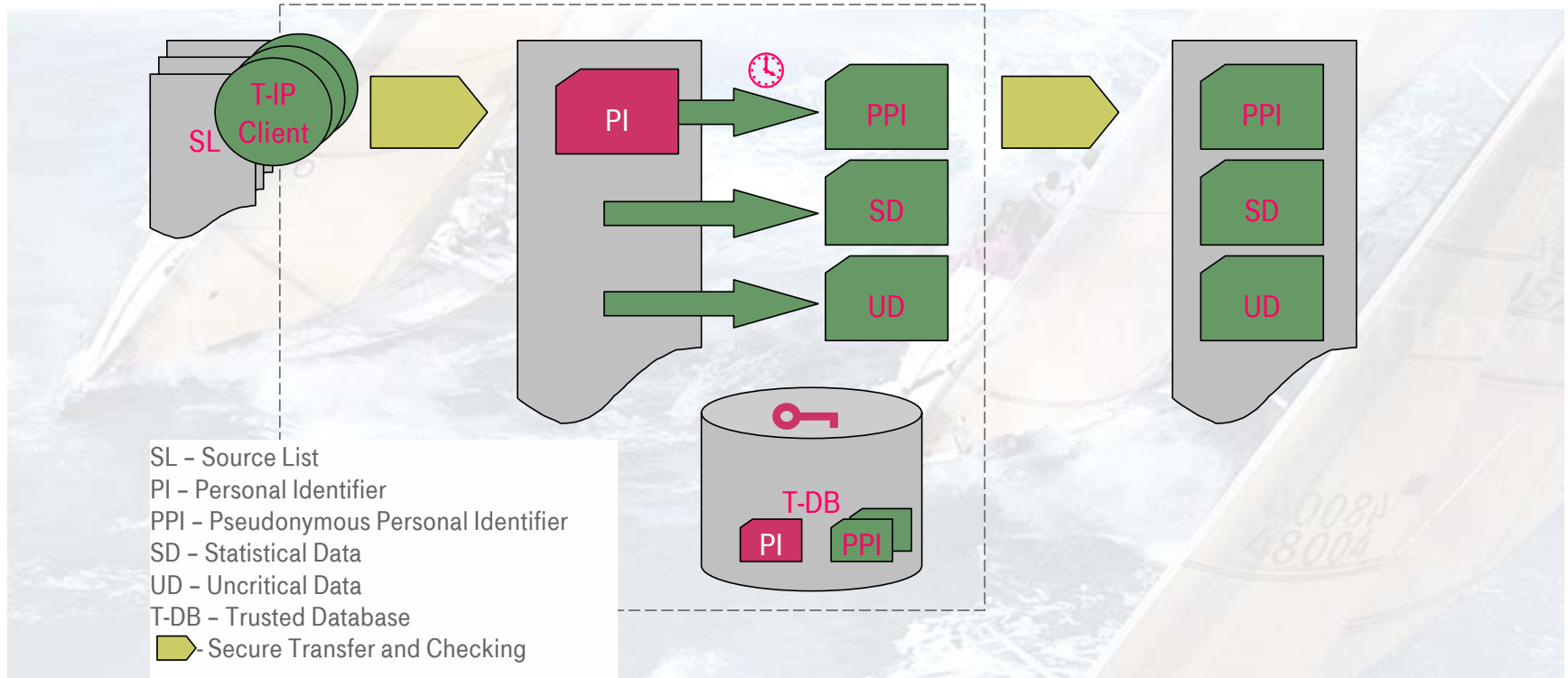
## Backup.

- More details of the T-IP



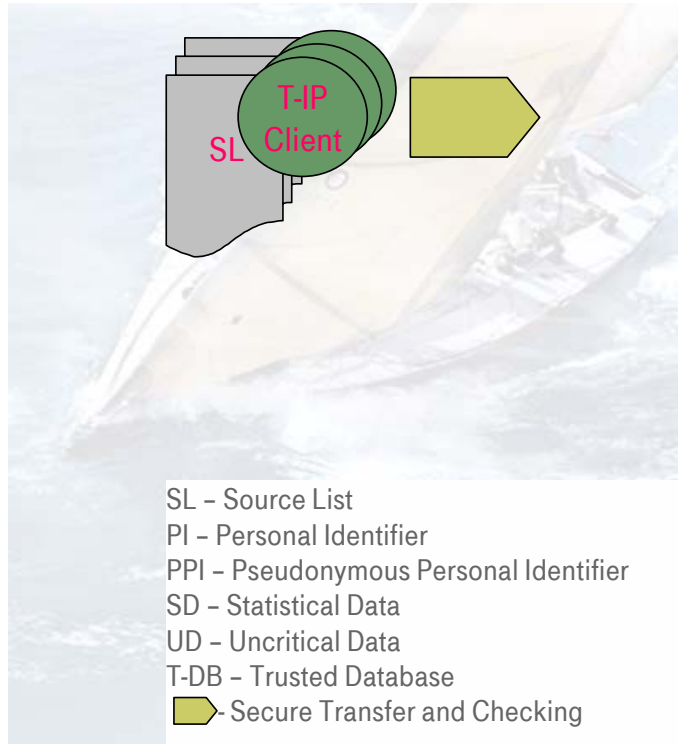
# T-Identity Protector.

## Methodology of pseudonymization – overview.



# T-Identity Protector.

## Methodology of pseudonymization – import.



### Data preparation in the source system (T-IP Client)

- Extraction of statistical characteristics from data fields that are pseudonymized.
- Pre-pseudonymization by converting the name and other identification characteristics into a pseudonym.
- Detection of erroneous data.

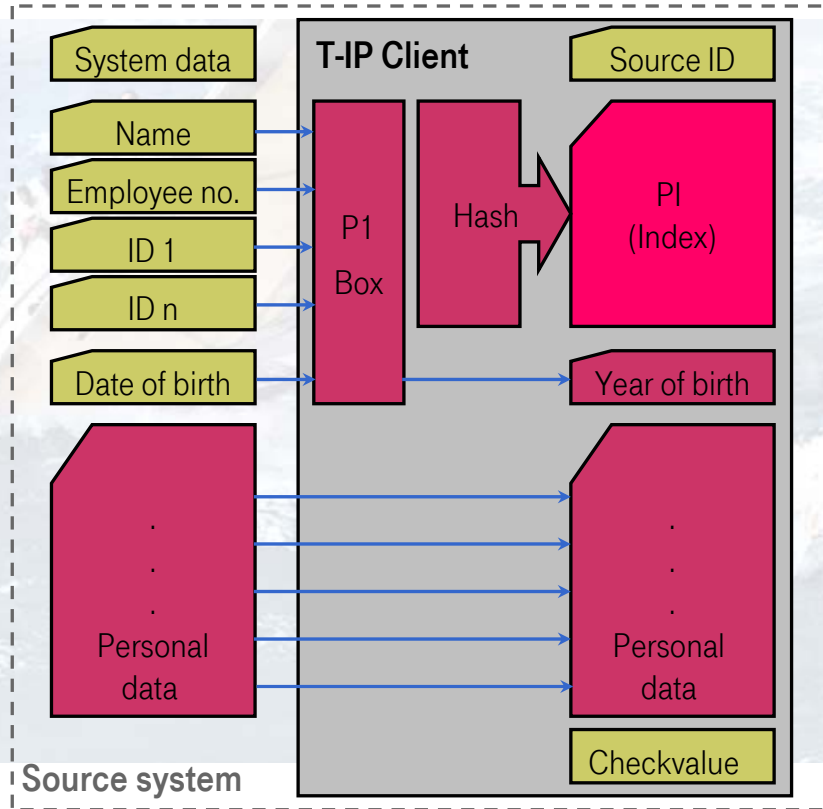
### Requirements for transfer

- Encrypted transfer.
- Reliable authentication with certificates.



## T-Identity Protector.

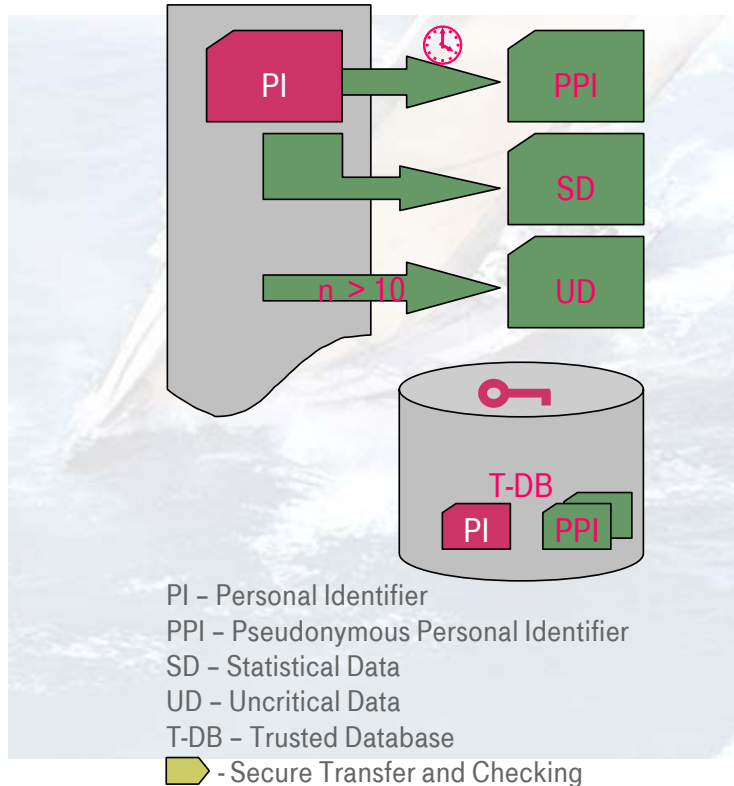
### Methodology of pseudonymization – pre-pseudonymization.



- The name and all data fields through which a person can be uniquely determined using other systems in the company are converted into a pseudonym.
- The P1 box protects against calculation of the hash value outside of the T-IP Client.
- The date of birth is reduced to the year of birth.
- The prepared data record is labeled with the source ID and secured with a check value for error detection.

## T-Identity Protector.

### Methodology of pseudonymization – data preparation.



#### Three data areas

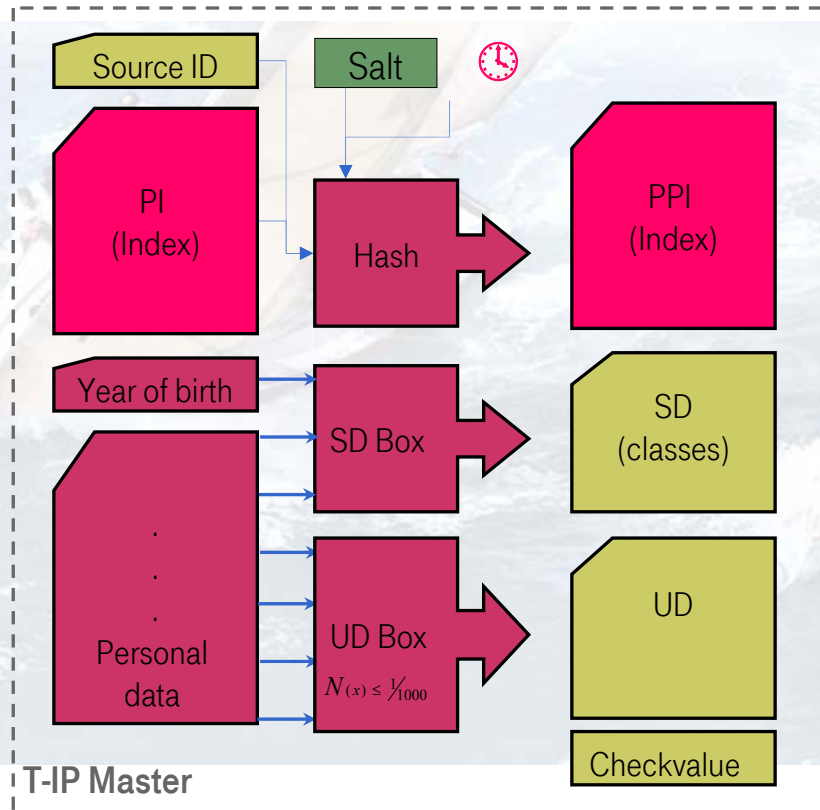
- Personal identification data with dependence on time (PPI)
- Statistical data in classes (SD)
- Uncritical data with a sufficiently high number of elements (DU)

#### Database controlled by the Trusted Agency

- Matching of source pseudonym with several time-dependent target pseudonyms.
- After a defined time or number of generations, target pseudonyms are deleted.

## T-Identity Protector.

Methodology of pseudonymization – pseudonymization.



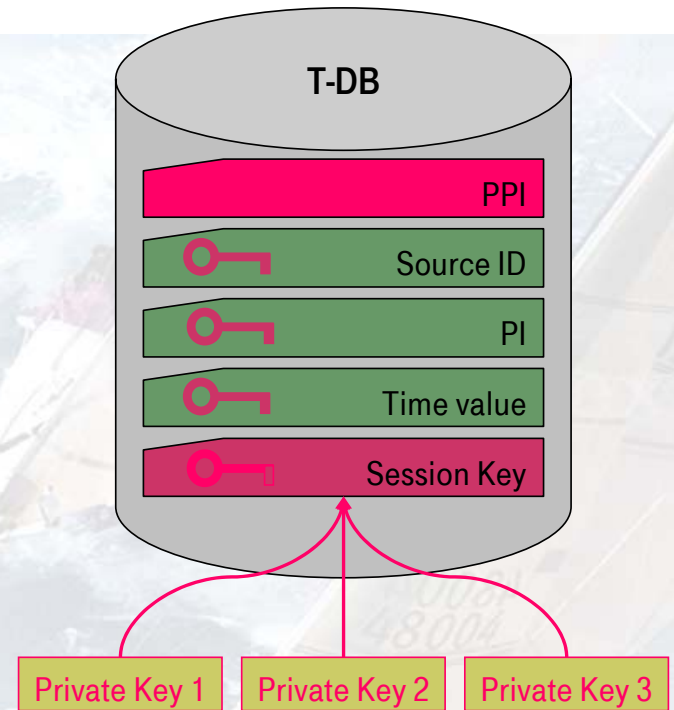
- The pseudonym (PPI) consists of the pre-pseudonym (PI), the source ID, a hash (Salt) and a time value.
- Statistical classes are set up in the SD box.
- Data fields with low frequency are overwritten in the UD box as “not analyzable.”



## T-Identity Protector.

### Methodology of pseudonymization – Trusted Database.

- The T-IP Master includes the Trusted Database (T-DB) which contains all data necessary for re-personalization.
- Source ID, PI and time value are encrypted symmetrically with a session key (AES).
- The session key is encrypted asymmetrically with the public key of the Trusted Agency and also stored in the database.
- The private key of the Trusted Agency is divided in a special process right when it is generated.
- The data can be used only if all parts of the private key are available (“knowledge distributed to different roles”).



Asymmetrical encryption (RSA)

## T-Identity Protector.

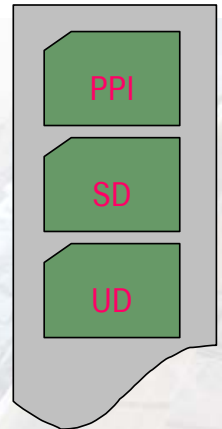
### Methodology of pseudonymization – export.

#### Requirements for transfer

- Encrypted transfer.
- Reliable authentication with certificates.

#### Pseudonymous data in the destination system


- The data can be analyzed in any manner in a data warehouse.
- Controlled re-personalization is possible only for a defined period of time.
- After the references in the Trusted Agency are deleted, this data can be used as anonymized data.



PPI – Pseudonymous Personal Identifier

SD – Statistical Data

UD – Uncritical Data

 - Secure Transfer and Checking

# T-Identity Protector.

## Methodology of pseudonymization – re-personalization.

