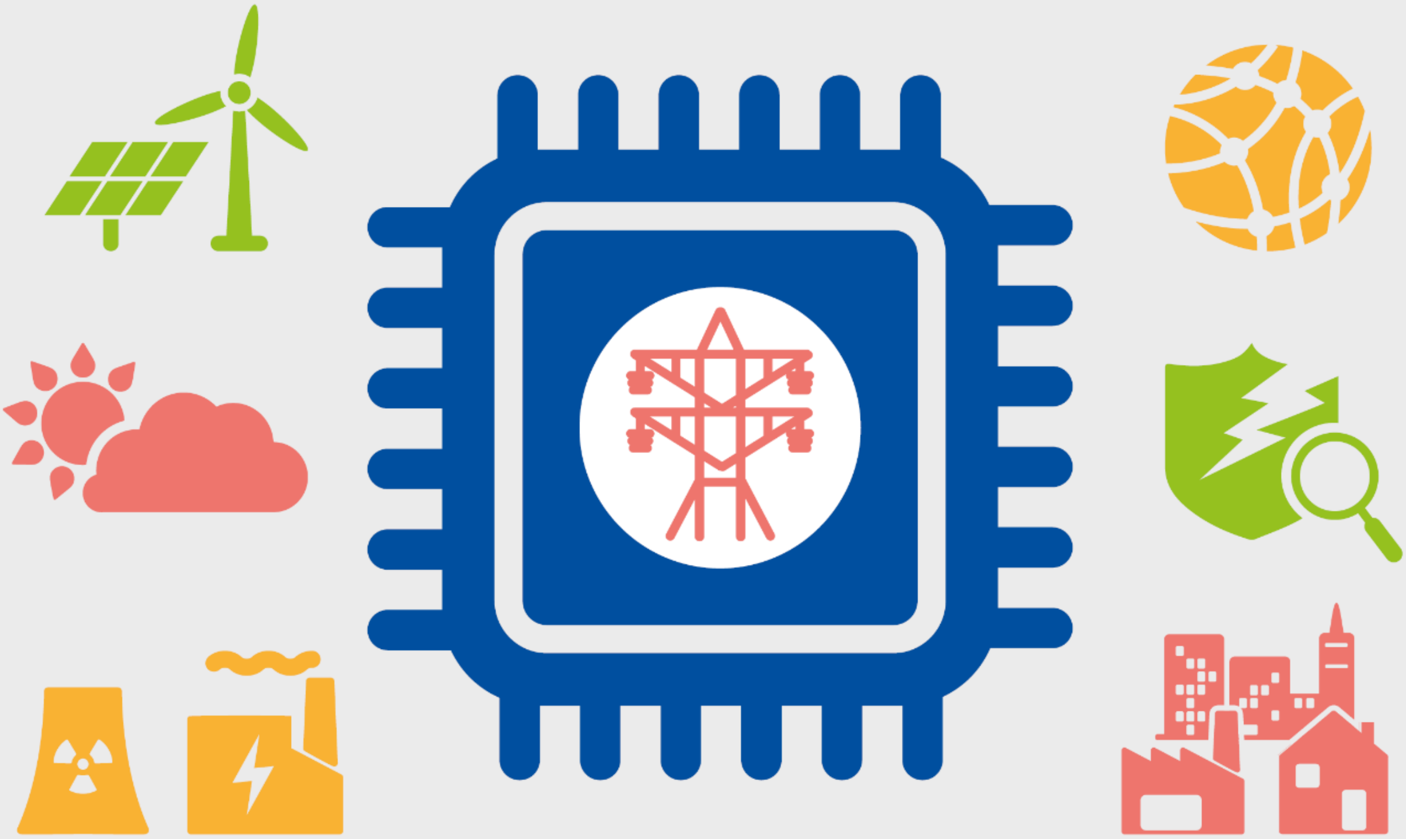
EUROPEAN UNION AGENCY
FOR CYBERSECURITY



CONTACT

EDITORS

LEGAL NOTICE

COPYRIGHT NOTICE

Cybersecurity Challenges of Artificial Intelligence[1] Algorithms[2]

Securing Machine Learning

This new report analyses cybersecurity and privacy requirements and measures in use of AI in forecasting demand on electricity grids. The report describes the scenario fundamental principles (assets, actors processes etc.), identifies the security and privacy risks it poses, and finally cybersecurity and privacy controls, which counteract the identified risks.

1.1 STUDY OBJECTIVES

- **Forecasting Demand on Electricity Grids**
-
-

1.2 METHODOLOGY

1.2.1 Description of the scenario

1.2.2 Identification of cybersecurity and privacy threats and vulnerabilities

1.2.3 Identification of cybersecurity and privacy controls

1.3 TARGET AUDIENCE

- **All actors (private or public):**

- **AI technical community, AI cybersecurity and privacy experts and AI experts**

- **Cybersecurity and privacy community**

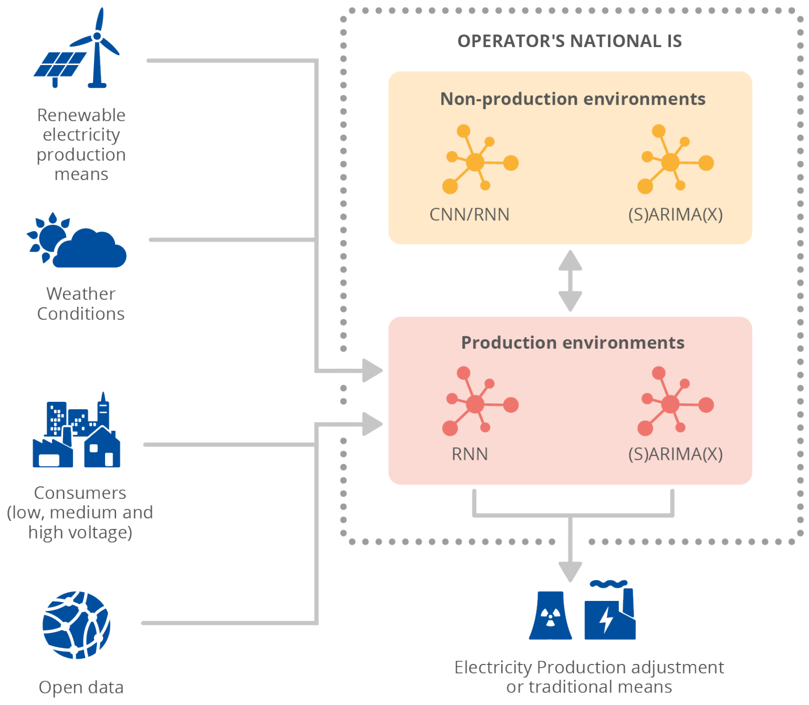
1.4 USING THIS DOCUMENT

Figure 1:

FORECAST DEMAND ON ELECTRICITY GRIDS

DESCRIPTION

An Electricity System Operator wants to prioritize, at aregional level, the conmsumption of electricity produced from renewable energies (e.g. photovoltaic panels, wind turbines). To do so, it uses machine learning to **forecast the daily production of electricity from renewable energies** and the **demand from consumers.** By comparing the two, he knows whether he needs to supplement (e.g. using conventional power plants) renewable energy production or not to meet the demand.


Renewable electricity production means
Weather Conditions
Consumers (low, medium and high voltage)
Open data
OPERATOR'S NATIONAL IS
Non-production environments
CNN/RNN
(S)ARIMA(X)
Production environments
RNN
(S)ARIMA(X)
Electricity Production adjustment or traditional means

DATA

Input data

- Material characteristics of renewable electricity production means
- High resolution weather data
- Demographics
- Current daily Energy consumption
- Calendar

Output data

- Energy production from renewable energies
- Energy consumption

CYBERSECURITY AND PRIVACY REQUIREMENTS

Cyber requirements

🟢 Availability 🔴 Integrity 🔴 Confidentiality 🟠 Traceability

Privacy Requirements

🟢 Availability 🟢 Integrity 🔴 Confidentiality 🟠 Traceability

- Lawfulness
- Fairness
- Transparency
- Purpose limitation,
- Data minimization,
- Accurancy
- Storage limitation
- Security of personal data
- Database creation,
- Compliance of the training model

🔴 Critical 🟠 High 🟢 Low

ACTORS

- Energy System Operator's teams
- Electricity consumers
- Open-data providers
- Data scientists
- Developers and Data Engineers
- System and communication network's administrator

ASSETS

- RNN & SARIMAX – renewable energy & consumption forecast algorithms
- Data lake & Model server – on premises
- Open Data provider APIs
- Smart Meter & Concentrator
- Operator's electrical grid
- Integrated Development Environment
- Libraries
- Communication protocols and network

2.1 PURPOSE AND CONTEXT

2.2 HIGH-LEVEL DESCRIPTION

supervised learning

collected in a data lake

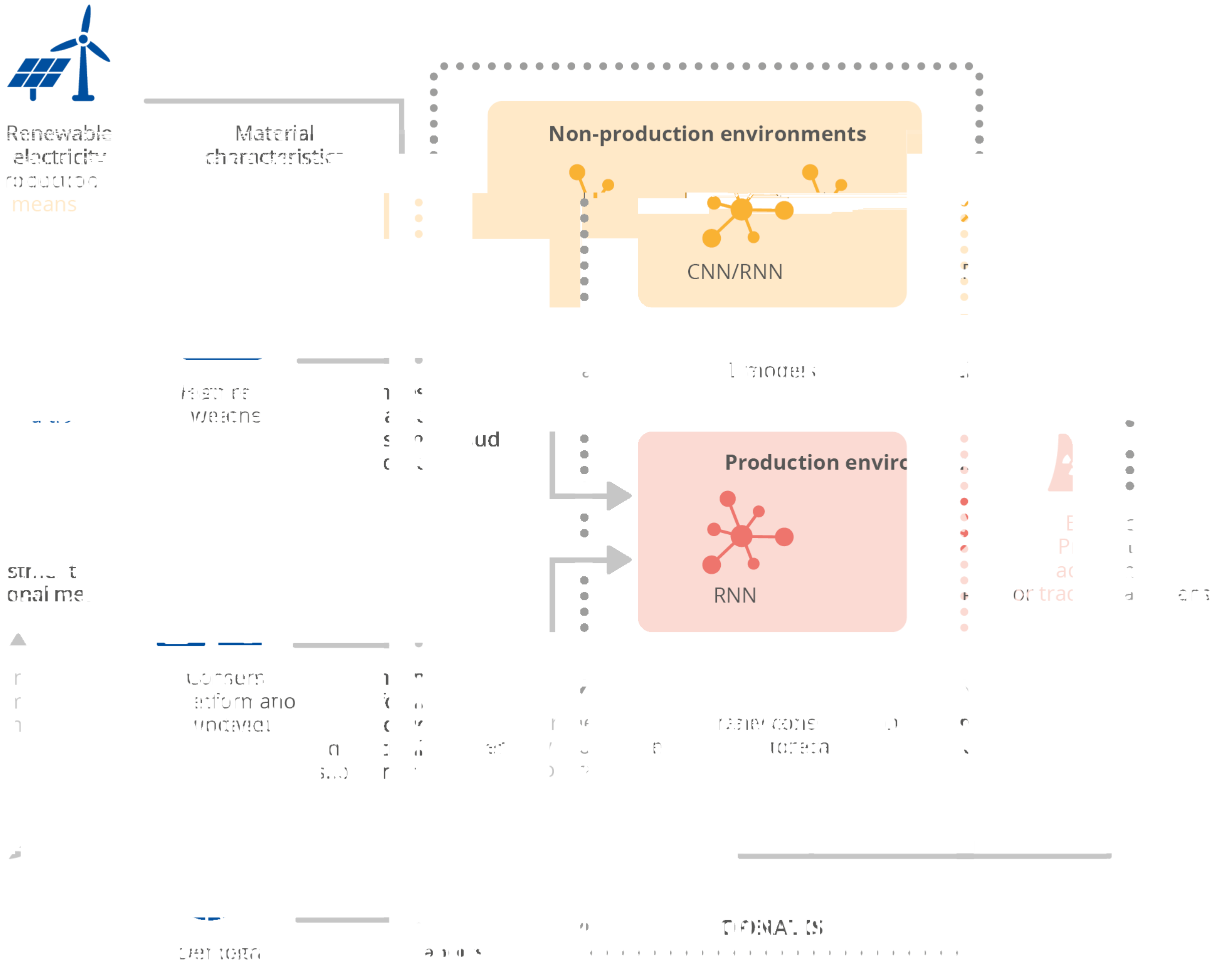
Selected data are then used to create two machine learning models.

These two models produce the following outputs

-
-

dashboard tool

Figure 2:



Renewable electricity production means
Material characteristics
Non-production environments
CNN/RNN
models
ud
Production enviro
RNN

2.3 ACTORS AND ROLES

Figure 3:

Actor	Role	Description
Electricity supplier's teams	*End Users and Data Owner (**Data Controller**)*	
Electricity consumers	*Data Provider*	
Open-data providers	*Data Provider*	
Data scientists	*Data scientist*	
Developers and Data Engineers	*Developers and Data Engineers*	
System and communication network administrators	*Network administrators*	

2.4 PROCESSED DATA

Figure 4:

Data	Data type	Source / data provider	Data Procurement
		the electricity system electricity supplier	
		Open-data provider	
		i.e., the electricity system electricity supplier	

Figure 5:

Data	Data type	Source / data provider	Data Procurement
		Open-data provider	
		Open-data provider	
		the electricity system electricity suppliers	

Elamin, Niematallah et Fukushige, Mototsugu.

		the electricity system electricity suppliers	

2.5 MACHINE LEARNING ALGORITHMS

Figure 6:

Learning paradigm	Subtype	Algorithm	Type of data ingested	Description

2.6 ASSETS

Neda Tavakoli; Sima Siami-Namini; Akbar Siami Namin.

B, Prabadevi, et al.

Abualig, Laith, et al.

Figure 7:

Type of asset	Asset	Description
Models		
Environment tools	*on-premises*	
	on-premises	
	on-premises	

2.7 OVERALL PROCESS

Data collection

(material characteristics, and production history
electricity supplier

weather
data

meteorological services
This data is considered as open-data and is therefore non-proprietary and free to use.

electricity consumption of the inhabitants

The electricity supplier collects their electricity consumption from smart meters installed locally at the consumers' place[21]

personal data

-
-
-
-

The default or detailed consumption values (if the consumer has agreed to share this information for the purpose of the processing) are kept and then aggregated (i.e., summed with all other consumption data) in a large consumption database

. This aggregated data does not allow for the retrieval of consumer data. It is therefore anonymised data.

demographics calendar

Data cleaning and data pre-processing

collected data cleaned

pre-processing

Model design and implementation

- A Recurrent Neuronal Network (RNN)
  - ○
- A Seasonal AutoRegressive Integrated Moving Average with eXogenous variables (SARIMAX)
  - ○

Brownlee, Jason.

Lee, Donghun and Kim, Kwanho.

Pavicevic, Milutin and Popovic, Tomo

Elamina, Niematallah et Fukushige, Mototsugu

Sim, Sze En, et al.

model's parameters

Model training, model testing and optimisation

training method

extreme weather conditions

Model Evaluation

evaluate the model

Model Deployment

model deployment

Monitoring and inference

Baheti, Pragati. *v7labs.*

monitoring

Figure 8:

Steps	Description	Actors	Assets
Data Collection			
Data Cleaning			
Data pre-processing			
Model design and implementation			
Model training			
Model testing			
Optimization			

Model evaluation			
Model deployment			
Monitoring and inference			

2.8 PRIVACY AND CYBERSECURITY REQUIREMENTS

Cybersecurity requirements

Figure 9:

	Level	Explanation
Availability	Low	every day half a week would be tolerable Longer unavailability
Integrity	Critical	accurate with a high level of quality large imbalance underproduction or an overproduction
Confidentiality	Critical	personal data (upstream of the concentrator)
Traceability	High	

Privacy requirements

It is important to note that the billing functions are not considered in our case, this topic being out of scope.

our scenario handles personal data in the data collection phase The following data protection requirements and recommendations should be satisfied

Figure 10:

Requirements	Explanation
Lawfulness, fairness, and transparency[30]	Lawfulness Fairness: Transparency:
Purpose limitation	
Data minimisation	

Accuracy	
Storage limitation	
Security of personal data (Integrity and Confidentiality)	

Figure 11:

Recommendations	Explanation
Database creation	
Compliance of the training model (i.e. before production)	

Figure 12:

Criteria	Does it match the criteria?	Justification

Figure 13:

	Level	Explanation
Availability	Low	
Integrity	Low	
Confidentiality	Critical	
Traceability	High	

3.1 THREAT CONTEXTUALISATION

electrical production disruption

reputation degradation,

phishing
attempts or targeted advertising

robbery

, separation/divorce

, or job loss

significant feeling of invasion of privacy, feeling out of
control of their personal data **change in energy consumption billing**

Figure 14:

COMPROMISE OF MACHINE LEARNING APPLICATION

PROD REP PHISH ROBB

- Poor access right management process
- Weak access control
- Use of vulnerable components

POISONING

PROD

- Lack of control poisoning
- Lack of data for increasing robustness to poisoning
- Use of unsafe data or model

HUMAN ERROR

REP INV PHISH ROBB

- Lack of documentation the Electrical forecast system
- Poor access rights management process
- Lack of security by design

DATA DISCLOSURE

INV PHISH ROBB

- Poor access right management process
- Weak access control
- Poor data management

UNLAWFUL PROCESSING

INV PHISH ROBB

- Lack of practical meansand justification for obtaining the consents of the electricity consumers concerned

UNFAIR PROCESSING

BILL INV PHI

- Absence of an identified data controller
- Lack of detail on the purposes and justification for their legitimacy
- Lack of traceability of actions and/or modifications made to the assets

LACK OF TRANSPARENCY

INV CONT

- Absence of an identified data controller
- Lack of justification for the collection of individual personal data collected
- Lack of transparency on the purpose of the use-case

DIVERSION OF PURPOSE

ROBB

- Lack of control of Data processor
- Lack of controls to ensure that data is used only for the purposes defined
- Lack of controls to ensure the adequacy of the purpose and its current use

NO RESPECT OF DATA MINIMIZATION

ROBB PHISH INV

- Lack of measures to prevent further data collection
- Lack of necessary data collection

NO RESPECT OF STORAGE LIMITATION

PHISH ROBB

- Lack of data deletion mechanisms
- Lack of data retention policy

IMPACTS

PROD:	Electrical production disruption	ROBB:	Robbery, separation/ divorce, job loss
REP:	Reputation degradation		
PHISH:	Phishing attempts, targeted advertising	INV:	Significant sense of invasion of privacy
BILL:	Change of consumer billing	CONT:	Not being in control of personal data

3.1.1 Compromise of ML application components

electrical production disruption

electrical production disruption

reputation degradation

significant feeling of invasion of privacy, phishing attempts, or targeted advertising

robbery, separation/divorce, or job loss

3.1.2 Poisoning

data collection

historical consumption data

electrical production disruption.

having collected the data

electrical production disruption

3.1.3 Human error

reputation degradation significant feeling of invasion of privacy, phishing attempts, or targeted advertising robbery, separation/divorce, or job loss

3.1.4 Data disclosure

reputation degradation. significant feeling of invasion of privacy, phishing attempts, targeted advertising, robbery, separation, or divorce and/or job loss.

3.1.5 Unlawful Processing

significant sense of invasion of privacy

a significant feeling of invasion of privacy

3.1.6 Unfair processing

unknowingly changing their billing

3.1.7 Lack of transparency

feeling of being not in control of personal data

3.1.8 Diversion of purpose

robbery, or separation/divorce, or job loss.

3.1.9 No respect of data minimisation

significant feeling of invasion of privacy **separation/divorce, job loss in case of data leakage** **potential phishing attempts, targeted advertising,** **robbery.**

3.1.10 No respect of storage limitation

separation/divorce, job loss, **phishing attempts, targeted advertising,** **robbery.**

3.1.11 Synthesis of possible impacts and associated threats

Figure 1:

Impact	Severity	Type	Associated Threats
Electrical production disruption	High		
Reputation degradation	High		
Phishing attempts, targeted advertising	Moderate		

Robbery, Separation/divorce or job loss	High		
Significant feeling of invasion of privacy	Moderate		
Not being in control of personal data	Moderate		
Change of consumer billing	High		

3.2 VULNERABILITIES ASSOCIATED TO THREATS AND AFFECTED ASSETS

Figure 2:

Vulnerabilities	Threats	Actors	Assets involved
Absence of an identified data controller			
Absence of mechanisms to ensure that processing of consumer electricity affected by consent cannot be carried out without consent			
Disclosure of sensitive data for ML algorithm training			
Existing biases in the ML model or in the data			
Lack of anonymisation			
Lack of auditability of processing			

Lack of control for poisoning			
Lack of control of Data processor[36]			
Lack of controls to ensure that data is used only for the purposes defined			
Lack of controls to ensure the adequacy of the purpose and its current use			
Lack of data deletion mechanisms			
Lack of data for increasing robustness to poisoning			
Lack of data retention policy			
Lack of detail on the purposes and justification for their legitimacy			
Lack of documentation			
Lack of justification for the collection of individual personal data collected			
Lack of legal basis related to users' consent when their detailed consumption data (per hour or half hour) are processed or that legitimate interest related to the daily processing of the data is not properly justified or that no justification is provided at all			

Lack of measures to prevent further data collection			
Lack of necessary data selection			
Lack of practical means and justification for obtaining the consents of the electricity consumers concerned (those who have a half-hourly view of their electricity consumption)			
Lack of security by design			
Lack of privacy by design			
Lack of security process to maintain a good security level of the components of the Electrical forecast system			
Lack of traceability of actions and/or modifications made to the assets			

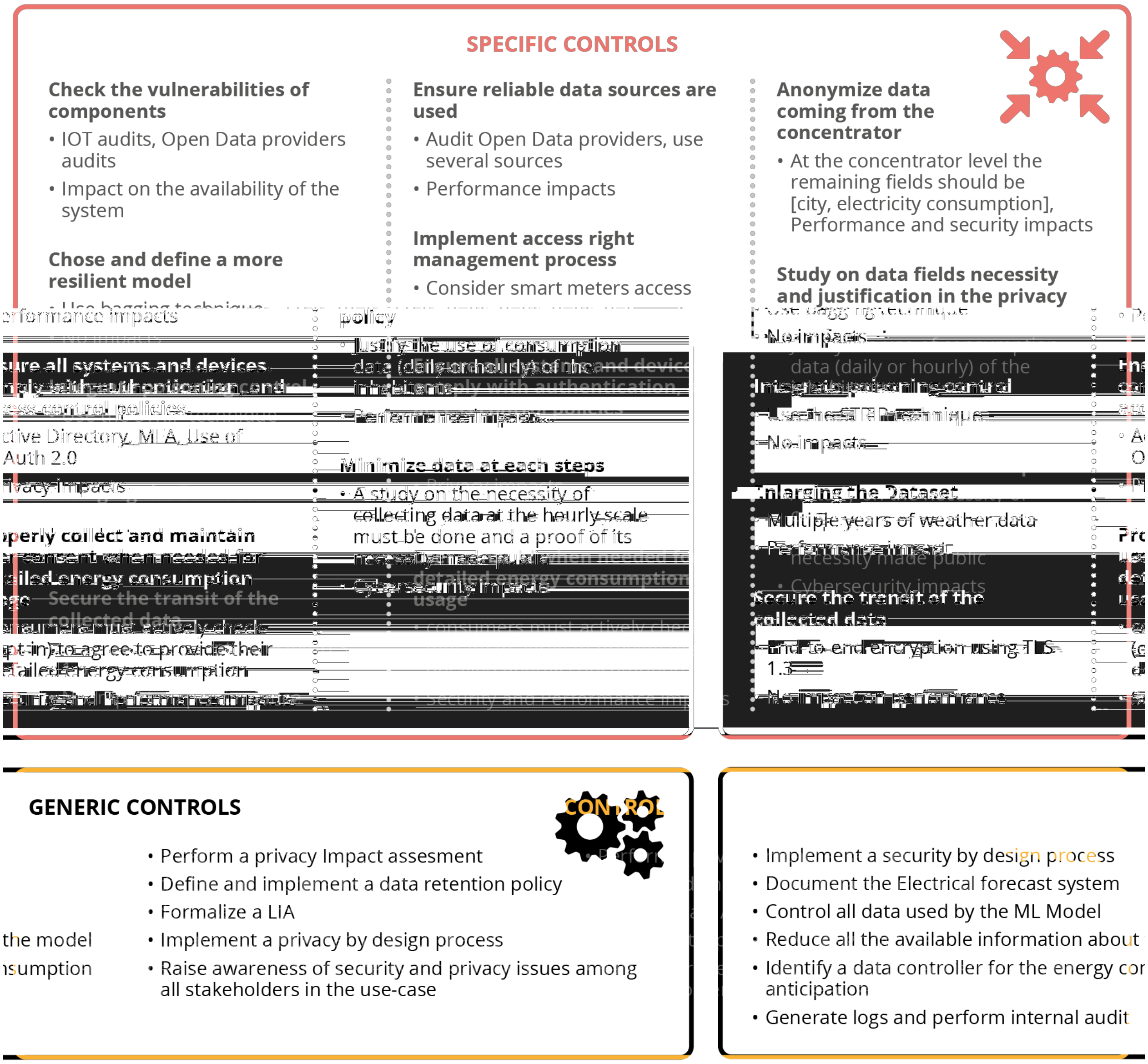
extracted, and how they are processed.			
Lack of verification that the data is adequate, relevant and not excessive for the purpose of estimating electricity consumption			
Model easy to poison			
No detection of poisoned samples in the training dataset			
Poor access rights management process			
Poor data management			
Excessive information available on the model			
Unprotected sensitive data on test environments			
Use of uncontrolled data			
Use of unsafe data or models (e.g., with transfer learning)			

Use of vulnerable components (Among the whole supply chain)			

Weak access protection

or.2 penael

Figure 17:



SPECIFIC CONTROLS

Check the vulnerabilities of components

- IOT audits, Open Data providers audits
- Impact on the availability of the system

Chose and define a more resilient model

- Use bagging technique

erformance impacts

ure all systems and devices ply with authentication and ess control policies

ctive Directory, MFA, Use of Auth 2.0

rivacy Impacts

perly collect and maintain r consent when needed for ailed energy consumption ge

onsumers must actively check pt-in) to agree to provide their tailed energy consumption

curity and Performance impacts

Ensure reliable data sources are used

- Audit Open Data providers, use several sources
- Performance impacts

Implement access right management process

- Consider smart meters access

policy

- Justify the use of consumption data (daily or hourly) of the inhabitants
- Performance impacts

Minimize data at each steps

- A study on the necessity of collecting data at the hourly scale must be done and a proof of its necessity made public
- Cybersecurity impacts

Anonymize data coming from the concentrator

- At the concentrator level the remaining fields should be [city, electricity consumption], Performance and security impacts

Study on data fields necessity and justification in the privacy

- Use bagging technique
- No impacts

Integrity poisoning control

- Use the STRIP technique
- No impacts

Enlarging the Dataset

- Multiple years of weather data
- Performance impact

Secure the transit of the collected data

- End-to-end encryption using TLS 1.3
- No impact on performance

GENERIC CONTROLS

the model

nsumption

- Perform a privacy Impact assesment
- Define and implement a data retention policy
- Formalize a LIA
- Implement a privacy by design process
- Raise awareness of security and privacy issues among all stakeholders in the use-case

- Implement a security by design process
- Document the Electrical forecast system
- Control all data used by the ML Model
- Reduce all the available information about
- Identify a data controller for the energy cor anticipation
- Generate logs and perform internal audit

4.1 IMPLEMENT A SECURITY BY DESIGN PROCESS

Type	Associated Vulnerabilities	Threats it mitigate
		• • • •

4.2 DOCUMENT THE ELECTRICAL FORECAST SYSTEM

Type	Associated Vulnerabilities	Threats it mitigate
		• • • • • • •

4.3 CHECK THE VULNERABILITIES OF THE ML COMPONENTS AND IMPLEMENT PROCESSES TO MAINTAIN THEIR SECURITY LEVELS OVER TIME

Type	Associated Vulnerabilities	Threats it mitigate
	• •	• •

4.4 CHOOSE AND DEFINE A MORE RESILIENT MODEL DESIGN

Type	Associated Vulnerabilities	Threats it mitigate

4.5 INTEGRATE POISONING CONTROL IN THE TRAINING DATASET

Type	Associated Vulnerabilities	Threats it mitigate
	• •	

4.6 ENLARGE THE TRAINING DATASET

Type	Associated Vulnerabilities	Threats it mitigate

4.7 SECURE THE TRANSIT OF THE COLLECTED DATA

Type	Associated Vulnerabilities	Threats it mitigate

4.8 CONTROL ALL DATA USED BY THE ML MODEL

Type	Associated Vulnerabilities	Threats it mitigate

4.9 ENSURE RELIABLE SOURCES ARE USED

Type	Associated Vulnerabilities	Threats it mitigate

4.10 IMPLEMENT ACCESS RIGHT MANAGEMENT PROCESS

Type	Associated Vulnerabilities	Threats it mitigate
		• • • •

4.11 ENSURE ALL SYSTEMS AND DEVICES COMPLY WITH AUTHENTICATION, AND ACCESS CONTROL POLICIES

Type	Associated Vulnerabilities	Threats it mitigate
	• •	• • • •

4.12 REDUCE THE AVAILABLE INFORMATION ABOUT THE MODEL

Type	Associated Vulnerabilities	Threats it mitigate

4.13 IDENTIFY A DATA CONTROLLER FOR THE ENERGY CONSUMPTION ANTICIPATION DATA PROCESSING

Type	Associated Vulnerabilities	Threats it mitigate
		• • •

4.14 PROPERLY COLLECT AND MAINTAIN USER CONSENT WHEN NEEDED FOR DETAILED ENERGY CONSUMPTION USAGE

Type	Associated Vulnerabilities	Threats it mitigate
	• • •	

4.15 ANONYMIZE DATA COMING FROM THE CONCENTRATOR

Type	Associated Vulnerabilities	Threats it mitigate

4.16 GENERATE LOGS AND PERFORM INTERNAL AUDIT

Type	Associated Vulnerabilities	Threats it mitigate
	• •	• • • • • • •

4.17 PERFORM A PRIVACY IMPACT ASSESSMENT

Type	Associated Vulnerabilities	Threats it mitigate

	• • •	• • • •

4.18 DEFINE AND IMPLEMENT A DATA RETENTION POLICY

Type	Associated Vulnerabilities	Threats it mitigate
	• •	

4.19 STUDY ON DATA FIELDS NECESSITY AND JUSTIFICATION IN THE PRIVACY POLICY

Type	Associated Vulnerabilities	Threats it mitigate
	• •	• • •

4.20 FORMALISE A LIA (LEGITIMATE INTEREST ASSESSMENT)

Type	Associated Vulnerabilities	Threats it mitigate

4.21 MINIMISE DATA AT EACH STEP OF THE PROCESSING; COLLECT ONLY WHAT IS NEEDED WHEN NEEDED

Type	Associated Vulnerabilities	Threats it mitigate
	• •	

4.22 IMPLEMENT A PRIVACY BY DESIGN PROCESS

Type	Associated Vulnerabilities	Threats it mitigate
		• • • • •

4.23 RAISE AWARENESS OF SECURITY AND PRIVACY ISSUES AMONG ALL STAKEHOLDERS

Type	Associated Vulnerabilities	Threats it mitigate
		• • • • • • •

4.24 SUMMARY

Figure 18:

Control name and type	Associated Vulnerabilities	Threat mitigated	Privacy and security requirements addressed
Implement a Security by Design process			
Document the Electrical forecast system			
Check the vulnerabilities of the components used and Implement processes to maintain security levels of ML components over time			
Choose and define a more resilient model design			
Integrate poisoning control in the training dataset			
Enlarge the training dataset			
Secure the transit of the collected data			
Control all data used by the ML Model			
Ensure reliable sources are used			

Implement access right management process			
Ensure all systems and devices comply with authentication, and access control policies			
Reduce the available information about the model			
Identify a data controller for the energy consumption anticipation data processing			
Properly collect and maintain user consent when needed for detailed energy consumption usage			
Anonymise data coming from the concentrator			
Generate Log generation and perform Internal audit process			
Perform a privacy Impact Assessment			

Define and implement a data retention policy			
Study on data fields necessity and justification in the privacy policy			

forecasting demand on electricity grids

A.1 CYBERSECURITY AND PRIVACY SEVERITY SCALES

Availability	
Low	few days or less
Moderate	a day or less
High	half a day or less
Critical	few hours or less

Integrity	
Low	does not need to be identified or corrected
Moderate	must be identified but not necessarily corrected
High	must be identified and corrected
Critical	No degradation

Confidentiality	
Low	accessed by everyone
Moderate	restricted to internal staff and trusted partners
High	restricted to employees having an organisation or functional link with the process
Critical	restricted to a very limited number of individuals

Traceability	
Low	absence of traces is acceptable
Moderate	Actions identified
High	actions actors identified and dated imputable
Critical	actions legally enforceable time stamped probative value

A.2 CYBERSECURITY SCALE OF IMPACT

Severity[42]	
1 - Low	
2 - Moderate	
3 - High	
4 - Critical	

A.3 PRIVACY SCALE OF IMPACT

Severity[43]	
1 - Low	
2 - Moderate	
3 - High	
4 - Critical	

A.4 PRIVACY REQUIREMENTS CRITERIA

Requirements	Article
Lawfulness, fairness and transparency	
Purpose limitation	
Data minimisation	
Accuracy	
Storage limitation	
Security of personal data (integrity and confidentiality)	

Recommendations	Details
Database creation	
Compliance of the training model (i.e. before production)	

ENISA
European Union Agency for Cybe

Athens Office
Agamemnonos 14, Chalandri 15

enisa.europa.eu