



**Digital
Water
.City**

D4.8 Cyber-physical systems protection training schemes

Including contributions to interoperability



Deliverable N° D4.8	Cyber-physical systems protection training schemes
Related Work Package	WP4
Deliverable lead	SINTEF
Author(s)	Martin Gilje Jaatun, SINTEF Audun Vennesland, SINTEF Rita Ugarelli, SINTEF Camillo Bosco, SINTEF Guillaume Bour, SINTEF
Contact for queries	martin.g.jaatun@sintef.no
Grant Agreement Number	n° 820954
Instrument	HORIZON 2020
Start date of the project	01 June 2019
Duration of the project	42 months
Website	www.digital-water.city
Abstract	This deliverable presents material for a training scheme on cyber-physical security for water network operators. It also reports on efforts on interoperability between DWC solutions and legacy systems, as well as documenting dissemination work on cyber-physical security training.

Dissemination level of the document

<input checked="" type="checkbox"/>	PU	Public
<input type="checkbox"/>	PP	Restricted to other programme participants
<input type="checkbox"/>	RE	Restricted to a group specified by the consortium
<input type="checkbox"/>	CO	Confidential, only for members of the consortium

Versioning and contribution history

Version*	Date	Modified by	Modification reasons
D1	2022-10-31	MGJ	Draft
R1	2022-11-23	MGJ	Draft following internal review (Sofia Housni, SIAAP and Adriano Mancini, UniVPM)
S	2022-11-24	Nico Caradot	Final review by coordinator
R2	2022-12-21	MGJ/RU	Update after EC review
V	2023-01-31	Nico Caradot	Final review before submission

* The version convention of the deliverables is described in the Project Management Handbook (D7.1). *D* for draft, *R* for draft following internal review, *S* for submitted to the EC and *V* for approved by the EC.

Table of contents

List of figures	6
List of tables	6
Glossary	7
Executive summary	8
1. Introduction	9
2. Contributions to interoperability	10
2.1. Definition of design requirements for interoperability	10
2.2. New data exchange standards in the Smart Data Models initiative.....	10
2.3. DWC Reference Ontology	10
2.4. Change requests for new classes and relationships in relevant SAREF ontologies	10
2.5. An architecture specification describing how interoperable data exchange can be accomplished.....	11
2.6. Implementation of middleware to ensure interoperable data exchange in the Paris case ..	11
2.7. Interaction with other research projects and initiatives	11
2.7.1. DigitalWater2020 / FIWARE/ICT4Water	11
2.7.2. ETSI SmartM2M	11
2.7.3. H2020 B-WaterSmart.....	12
2.7.4. Policy recommendations	12
3. Training schemes for cyber-physical security	13
3.1. Course Introduction – introduction to CP security	13
3.2. Water systems as cyber-physical entities	13
3.3. Information Systems for water	14
3.4. Communication technologies for water	14
3.5. Risk Management for Cyber-physical Security	14
3.6. IoT Security for water systems.....	15
3.7. Organizational Resilience Training.....	15
3.8. Course end	15
4. List of dissemination efforts.....	16
4.1. Dissemination details	17
4.2. Acknowledgment of training material in the water community	17
4.3. Awarding of diploma or other means of formal recognition.....	17
5. Conclusion	18
6. References.....	18
Appendix A: Sample chapters.....	23
A.1. Sample material for chapter: Introduction	23
Topic.....	23
Goals.....	23
Text.....	23

Other industries	24
Fundamental problems	24
Structure of the training material	24
A.2. Sample material for chapter: Water systems as cyber-physical entities.....	26
Topic	26
Goals.....	26
Instructor pre-requisites	26
Text.....	26
EXAMPLE of relevant cyber attack.....	29
Modeling approaches.....	30
Evaluation questions	32
A.3. Sample material for chapter: Information systems for water	35
Topic	35
Goals.....	35
Instructor pre-requisites	35
Text.....	35
Introduction	35
Supervisory Control and Data Acquisition	35
Evaluation questions	38
A.4. Sample material for chapter: Communication technologies for water	40
Topic	40
Goals.....	40
Instructor pre-requisites	40
Text.....	40
Evaluation questions	45
A.5. Sample material for: Risk management process for cyber-physical security	47
Topic	47
Goals.....	47
Instructor pre-requisites	47
Text.....	47
The risk management process described by the ISO 31000:2018 standard.....	47
Cyber-physical risk assessment tools for the water sector	53
The risk management process applied to the water sector protection against C-P threats	57
A.6. Sample material for chapter: IoT security for water	70
Topic	70
Goals.....	70
Instructor pre-requisites	70
Text.....	70
Introduction & Background.....	70

Threat Landscape for Digital Solutions 74

Security Testing Methodology for IoT Devices 77

Best practices for IoT-based Solution Development..... 79

Common attacks and security vulnerabilities on sensors/IoT 84

A.7. Sample material for chapter: Scenario-based training exercise..... 86

Topic..... 86

Goals..... 86

Instructor pre-requisites 86

Text..... 86

 Overview of TORC 86

 Playing the TORC game 87

 STOP-IT Case: denial of service due to signal jamming..... 93

 DWC Case: spoofing of a web application impacting the service 94

List of figures

Figure 1: Example diploma after finishing course	16
Figure 2: Overview of the interactions of cyber and physical elements in a CPS	28
Figure 3: Example of the RISKNOUGHT simulation-based tool for CPS and the different elements modeled with it.	31
Figure 4: PLC from Siemens.....	36
Figure 5: Typical infrastructure of a SCADA system	37
Figure 6: TCP/IP protocol stack	41
Figure 7: Purdue model.....	44
Figure 8: The risk management process described by the ISO 31000	48
Figure 9: Example of risk matrix.....	50
Figure 10: Use of risk matrix in the risk evaluation step: numbers 1-7 refer to analyzed risk events..	51
Figure 11: Lowering risk levels through adoption of RRM.....	52
Figure 12: Records structure in the RIDB	53
Figure 13: Selected event following the generic records structure of RIDB	59
Figure 14: Measured flows at the entrance of the analysed Biofos WWTP in the year 2020	60
Figure 15: Critical events in 2020 identified with the stress-testing procedure.....	61
Figure 16: Rain data in 2020 of seven stations of the Danish rain gauge network.....	62
Figure 17: Risk reduction measures in the RRMD associated with the identified risk in the RIDB.....	68
Figure 18: Critical hourly values in 2020 of wastewater without any rain contribution in the case of doubling the volume of the equalization tanks as risk reduction measure	69
Figure 19 Generic Architecture Diagram for a Digital Solution (in DWC)	74
Figure 20: Classification of the attacks against Digital Solutions (in DWC).....	76
Figure 21: High level diagram of the Black Box Testing Methodology.....	78
Figure 22: Key features of playing the TORC game	88

List of tables

Table 1: Cyber-physical security dissemination efforts.....	16
Table 2: Characterization of the considered water system	57
Table 3: Pre-defined level of Risk expressed in terms of the selected KPI.....	58
Table 4: Identification of the level of risk by comparing results with targets values.....	67
Table 5: TORC quick-start guide (illustrations derived from STOP-IT project [28])	90
Table 6: Example of playing with STOP-IT TORC	93
Table 7: Example of playing with TORC for a DWC case	94

Glossary

ETSI	European Telecommunications Standards Institute
FIWARE	A curated framework of OpenSource Platform components to accelerate the development of Smart Solutions (https://www.fiware.org)
SAREF	Smart Applications REference ontology (https://saref.etsi.org/)
STOP-IT	Strategic, Tactical, Operational Protection of water Infrastructure against cyber-physical Threats (https://stop-it-project.eu)

Executive summary

This deliverable presents material for a training scheme on cyber-physical security for water network operators.

The training material comprises the following modules:

- Course Introduction – introduction to CP security
- Water systems as cyber-physical entities
- Information Systems for water
- Communication technologies for water
- Risk Management for Cyber-physical Security
- IoT Security for water systems
- Organizational Resilience Training

This deliverable also reports on efforts on interoperability between DWC solutions and the link with different water ontology consortia, as well as documenting dissemination work on cyber-physical security training.

1. Introduction

Cybersecurity maturity in the water sector is growing, but there is a gap between large and medium-small water utilities. Existing solutions applied by large utilities might not be well known in the sector and may not be easily transferable or down-scalable to Small and Medium-sized Utilities (SMU). SMU also lack the personal/technical and financial resources to test and adapt promising solutions, at Technology Readiness Level (TRL) below market-readiness, on their own initiative, without scientific and financial support.

Also, the introduction of new digital systems and devices in the operation of water systems requires new types of expertise: water organizations should heavily invest in security education and training, as well as in IT security awareness campaigns.

This document describes a proposal for a training scheme for increasing competence in cyber-physical security for the water and wastewater sector. The document also contains a description of contributions to interoperability from the DWC project.

The remainder of this document is structured as follows:

Section 2 summarizes WP4's contributions to interoperability, both within the project and in relation to other projects and water ontology consortia.

Section 3 provides an overview of the training material created by and/or further developed by DWC.

Section 4 enumerates the various outreach activities with respect to cyber-physical security training performed by WP4.

Section 5 concludes the deliverable.

Appendix A contains sample material that can be used as a foundation for a cyber-physical training scheme.

2. Contributions to interoperability

Interoperability is a challenge in the water management domain, as in most other domains. Challenges relate to, among other things, legacy systems using proprietary data formats, closed or sparse interface definitions, and lack of data exchange standards. Deliverables D4.4 [50], D4.5 [56], and D4.7 [47] report how work package 4 have approached these challenges from different perspectives and made the following contributions to interoperability:

2.1. Definition of design requirements for interoperability

Building on initial system and data requirements collected from interviews with water utilities and digital solution providers participating in the project (reported in D4.1 [46]), we defined a set of design requirements relevant for proposing a semantic interoperability middleware architecture in the DWC project. These design requirements were defined in collaboration with representatives from the Paris and Milan use cases of DWC. Once the design requirements were mature enough, they were mapped with existing standard models developed by the Smart Data Models initiative, the SAREF¹ ontologies developed by ETSI, and the data exchange infrastructure components provided by FIWARE. The final set of design requirements, which are reported in D4.4 [50], guided further developments of standardised data models in the Smart Data Models initiative, the DWC Reference Ontology and the interoperability middleware.

2.2. New data exchange standards in the Smart Data Models initiative

Based on the mapping between the defined design requirements and existing standard data models in Smart Data Models², new data models and attributes were proposed. This work was performed in collaboration with FIWARE, one of the founding partners of the Smart Data Models initiative. A summary of the standardisation approach and concrete contributions made from DWC to Smart Data Models is reported in deliverable D4.7 [47].

2.3. DWC Reference Ontology

The DWC Reference Ontology, described in detail in D4.5 [56], is an OWL³ ontology that brings in concepts from existing SAREF ontology as well as other ontologies, e.g. for expressing temporal and spatial concepts. The DWC Reference Ontology specifies a semantic description of water management concepts through a structuring, classification, and formal definition of these concepts. The ontology can be used to mediate between heterogeneous formats (e.g.,[48]), as a contextual reference using the data exchange standards promoted by Smart Data Models initiative, and to support data analytics, for example using graph representation learning techniques (e.g.,[49]).

2.4. Change requests for new classes and relationships in relevant SAREF ontologies

The mapping between design requirements and existing semantic models also identified the need for additional classes and relationships in the SAREF ontologies relevant for DWC. This led to the submission of change requests to the ETSI Smart M2M technical committee responsible for

¹ <https://labs.etsi.org/rep/saref>

² <https://smartdatamodels.org/>

³ Web Ontology Language

maintaining the ontologies. A summary of the submitted change requests along with an explanation of the required procedures is reported in deliverable D4.7 [47].

2.5. An architecture specification describing how interoperable data exchange can be accomplished

In collaboration with representatives from the Paris and Milan cases of DWC a realisation architecture was developed to specify how interoperable data exchange could be implemented. The realisation architecture included relevant FIWARE components (e.g., a context broker, relevant database solutions for persisting water management entities, and IoT agents) and necessary middleware components needed to transform data between legacy systems and the FIWARE infrastructure. This reference architecture was used as a blueprint in the Paris implementation. The architecture is reported in deliverable D4.4 [50].

2.6. Implementation of middleware to ensure interoperable data exchange in the Paris case

To further support interoperable data exchange, the project developed a middleware component that transforms between proprietary data formats and data formats standardised in Smart Data Models initiative. Based on design requirements and the reference architecture reported in D4.4 [50] the middleware system specifically transforms data described using a proprietary format from the WWTP system in Paris to a FIWARE Context Broker hosted by KWB. The FIWARE Context Broker hosts water management data represented according to data models from the Smart Data Models initiative and are used as input to the water quality predictions performed by KWB. Due to time and resource constraints, it was not possible to implement the middleware in the Milan case. The middleware is reported in deliverable D4.7 [47].

2.7. Interaction with other research projects and initiatives

Other ongoing projects are also dealing with interoperability challenges in the water management domain, as detailed below.

2.7.1. DigitalWater2020 / FIWARE/ICT4Water

DWC has established strong links with the DigitalWater2020 group and ICT4Water cluster (action group “Enable Data Sharing”) mainly through the direct collaboration with FIWARE and specifically with Alberto Abella (FIWARE). Meetings with him related to including data elements defined in DWC into relevant data models standardised in smartdatamodels.org. This collaboration resulted in a new data model called WaterQualityPredicted, as well as extensions added to the existing data models Device, WaterQualityObserved, WaterObserved, and AgriParcel. A summary of the standardisation approach and concrete contributions made from DWC to Smart Data Models is reported in deliverable D4.7 [47].

2.7.2. ETSI SmartM2M

Contributing to future versions of SAREF ontologies based on ontology development reported in D4.5 and D4.7 [47]. The contributions are expressed as change requests to the public forge⁴ maintained by

⁴ <https://labs.etsi.org/rep/saref/saref-core/-/issues>

ETSI. The change requests include new ontology concepts and properties in the SAREF, SAREF4WATR and SAREF4AGRI ontologies maintained by ETSI technical committees.

2.7.3. H2020 B-WaterSmart

As in DWC, the B-WaterSmart project⁵ is also using the combination of FIWARE, Smart Data Models and SAREF ontologies as means to support data exchange within the domain, and lessons learned from DWC have been shared with B-WaterSmart [51]. In our collaboration with the B-WaterSmart project we shared experiences from requirements collection in DWC as input to requirements collection in B-WaterSmart. Furthermore, ideas and experiences from technical implementations of interoperability middleware (D4.7 [47]) were discussed with representatives from B-WaterSmart.

2.7.4. Policy recommendations

The work performed in DWC in rising awareness in the water sector, both in terms of semantic interoperability and cybersecurity, has been also documented as recommendations for Policy Developments at EU Level [62].

⁵ <https://b-watersmart.eu/>

3. Training schemes for cyber-physical security

The training schemes which will be presented in this deliverable build on the results on cyber-physical security in DWC water value chains [27][57], extending previous efforts of the H2020 STOP-IT project [5]. The aim is to create a valuable training material package to support accreditation and training schemes on cybersecurity tailored for the water sector.

The H2020 STOP-IT project produced a proposal for a training scheme for cyber-physical security [5]. This scheme was divided into a basic and an advanced part, where the advanced part centred around the specific tools developed in STOP-IT. For DWC, however, the most re-use potential can be found in the basic part, which is independent of proprietary technology.

Starting from the basic part, the material developed in STOP-IT is enriched by the knowledge and the products developed by T4.2, such as the risk management guide, the RIDB, the RRMD and the check list of IoT security for water systems.

In the following, we will sketch the content of a cyber-physical (CP) security training scheme based in part on STOP-IT [5], which then will be elaborated in later sections.

3.1. Course Introduction – introduction to CP security

This module introduces the material and outlines the content of the training package.

Summarized:

- introduction to cyber-physical security;
- some motivational examples from the past;
- outline of the course material;
- videos [STOP-IT introduction video, 2'18"].

See sample content in Appendix A.1

3.2. Water systems as cyber-physical entities

It introduces how and why utilities are going digital, in what ways (sensors, etc.) and how that means more CP risk, and offers a basic overview of how a water utility functions as a cyber-physical system (two interconnected layers, sensor and actuator principle). Further examples of CPS attacks are provided: cyber-attacks, physical impacts (and vice versa).

Summarized:

- an intro to how and why utilities are going digital, in what ways (sensors, etc.) and how that means more CP risk.
- basic overview of how a water utility functions as a cyber-physical system (two interconnected layers, sensor and actuator principle).
- examples of CPS attacks: cyber attacks, physical impacts (and vice versa).
- chapter evaluation/quiz

See sample content in Appendix A.2

3.3. Information Systems for water

This module presents key information systems for water, with a special focus on SCADA units. Covers attack vectors relevant to these information systems (including physical, cyber and cyber-physical attacks).

Summarized:

- key information systems for water, with a special focus on SCADA units
- attack vectors relevant to these information systems (physical, cyber and cyber-physical attacks)
- chapter evaluation/quiz

See sample content in Appendix A.3

3.4. Communication technologies for water

Presents key communication technologies for (smart) water, wired and/or wireless, and the specific attack vectors relevant to these communication technologies (as above).

Summarized:

- key communication technologies for (smart) water, wired and/or wireless.
- attack vectors relevant to these communication technologies (physical, cyber and cyber-physical attacks)
- chapter evaluation/quiz

See sample content in Appendix A.4

3.5. Risk Management for Cyber-physical Security

Introduces the risk management process described by the ISO 31000:2018 standard and shows how it relates to cyber-physical risk assessment management. Introduces a cyber-physical risk assessment tool for the water sector.

Summarized:

- The risk management process described by the ISO 31000-2018
- Physical risk assessment management based on ISO 31000:2018
- Introduction to a cyber-physical risk assessment tool for the water sector
- STOP-IT Risk Identification DataBase (RIDB) and Risk Reduction Measure Database (RRMD) for cyber-physical risk assessment
- Guide for risk management
- Quiz

See sample content in Appendix A.5

3.6. IoT Security for water systems

Introduces the concept of IoT security for water systems and presents a concrete methodology for security testing of IoT devices, providing some "dos and don'ts" and good practice in form of a checklist.

Summarized:

- Introduction & Background
 - Shift to Industry 4.0 and the "fusion" of the IT and OT worlds
 - General background on IoT and embedded systems (skipped in sample)
 - Common network protocols for IoT (skipped in sample)
- Threat landscape to IoT for water systems
 - Type of attacks against digital solutions
 - Classes of attackers
- Methodology for security testing of IoT devices
- Best practices

See sample content in Appendix A.6

3.7. Organizational Resilience Training

Introduces the online Training for Operational Resilience Capabilities (TORC) game, and documents how to play TORC as a teaching experience.

Summarized:

- Introduction
- Overview of the Training for Operational Resilience Capabilities (TORC) game⁶
- Play TORC as teaching experience
 - STOP-IT case
 - DWC case

See sample content in Appendix 0

3.8. Course end

If all evaluated chapters are finished successfully (Successful questions above a set threshold, e.g. >70%), the user passes the course and can be awarded a diploma.

⁶ <http://www.torc.no>

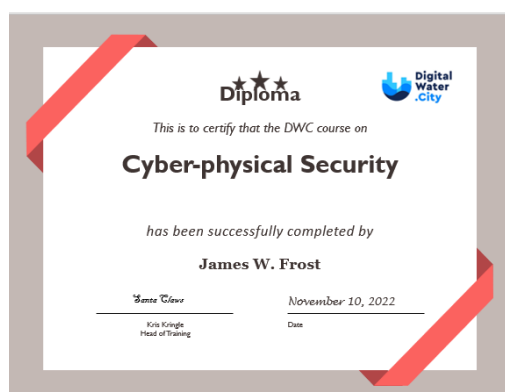


Figure 1: Example diploma after finishing course

4. List of dissemination efforts

Participants in T4.4 have performed several activities in order to promote dissemination of security knowledge for cyber-physical systems. These activities are mentioned briefly in Table 1. Further information on future opportunities for exploitation are provided in the following sections.

Table 1: Cyber-physical security dissemination efforts

When	What	Who
2021-2022	Discussions with the Erasmus+ project Digiwater regarding the possible inclusion of material on cyber-physical security in Digiwater university course: http://waterharmony.net/digiwater/	Rita Ugarelli Martin Gilje Jaatun
2021-2022	Discussions with European stakeholders (ENISA) on establishment of the European Water ISAC	Rita Ugarelli Martin Gilje Jaatun
January 18 2022	Presentation of training scheme for potential members of the Water ISAC	Martin Gilje Jaatun
May 2022	Lecture for TU Berlin	Rita Ugarelli
June 23 2022	Presentation of security assessment of a water operator IoT device, EU Water ISAC meeting	Guillaume Bour
September 11-15	Presentation of DWC poster and pitch at IWA WWC	Camillo Bosco
September 28-30 2022	TU Berlin workshop	Rita Ugarelli Martin Gilje Jaatun
October 14 2022	Communities of Practice meeting on cyber-physical security (online)	Camillo Bosco Guillaume Bour Martin Gilje Jaatun Gema Raspati
November 2022	Dissemination campaign	ARTICK

4.1. Dissemination details

The training material included in this deliverable has been disseminated in various stages as detailed in Table 1. Moving forward, the dissemination campaign, performed by ARTICK, has created a "teaser" consisting of the first module ("Water systems as cyber-physical entities"), which has been uploaded to the project website⁷ pending the availability of the full deliverable. Furthermore, a short video⁸ raising awareness on the topic and inviting to download the training material, has been created and been disseminated via social media.

4.2. Acknowledgment of training material in the water community

In collaboration with ENISA and the Empowering EU ISACs project⁹ we have been exploring possibilities for establishing a European-wide Information Sharing Analysis Centre (ISAC) for the water sector. As part of the preparations for starting a European Water ISAC, we have organized a number of webinars with participation of several European water utilities, and there presented the overview of the DWC training modules. The feedback from the utilities was overwhelmingly positive, and they welcomed the prospect of having training in cyber-physical security as part of possible future European Water ISAC events.

We have also organized two Communities of Practice (CoP) events within the DWC project, where different parts of the training material have been presented and well received by the representatives from the water community.

This work will continue as part of the Horizon Europe CSA European Knowledge Hub and Policy Testbed for Critical Infrastructure Protection (EU-CIP) project¹⁰.

4.3. Awarding of diploma or other means of formal recognition

The diploma in Figure 1 is intended as an example. DWC has had discussions with the Digiwater Erasmus project, and made the DWC training material available to them for possible further adoption at partner universities in the future. We have had more concrete discussions with the Norwegian University of Science and Technology (NTNU) and Technische Universität Berlin, and have agreed to use selected modules of the DWC training material in pilot courses at both universities in the spring semester of 2023. Course completion certificates will be issued by the individual university.

Further opportunities for exploitation of the DWC training modules exist as part of the Horizon Europe CSA EU-CIP (where at least one of the modules will be used), and possibly the Horizon Europe RIA *Ready!* (submitted under the HORIZON-CL3-2022-DRS-01-02 call, evaluation pending).

⁷ <https://www.digital-water.city/wp-content/uploads/2022/12/Water-systems-as-cyber-physical-entities.pdf>

⁸ <https://www.digital-water.city/news/water-systems-cyber-physical-entities-trainin-secure-system/>

⁹ <https://www.isacs.eu/>

¹⁰ <https://www.eucip.eu/>

5. Conclusion

This deliverable has presented material for cyber-physical security training for water network operators that can be used as a basis for an industry training program.

The deliverable has also documented work on interoperability between DWC solutions and other initiatives in the water domain.

6. References

- [1] Bernsmed, Karin, Per Håkon Meland, and Martin Gilje Jaatun, "Cloud Security Requirements - A checklist with security and privacy requirements for public cloud services." (2015)
- [2] Jaatun, Martin Gilje, et al. "Security checklists: a compliance alibi, or a useful tool for water network operators?" *Procedia Engineering* 70 (2014): 872-876.
- [3] Jaatun, Martin Gilje, Jon Røstum, and Stig Petersen, "Sikkerhet og sårbarhet i driftskontrollsystemer for VA-anlegg (in Norwegian)," Norsk Vann R195, 11. mars 2013, <http://www.norskvann.no/kompetanse/va-bokhandelen/rapporter/product/418-r195-sikkerhetog-sarbarhet-i-driftskontrollsystemer-for-vaanlegg>
- [4] Røstum, Jon and Martin Gilje Jaatun, "Informasjonssikkerhet og skybaserte tjenester for vannbransjen (in Norwegian)," May 2018, Norsk Vann rapport 238/2018. <https://norskvann.no/index.php/kompetanse/va-bokhandelen/produkt/681-a238-informasjonssikkerhet-og-skybaserte-tjenester-for-vannbransjen>
- [5] Bouziotas, Dimitrios and Lisa Andrews, "Continuity plan for the training courses and services", STOP-IT Deliverable D8.4, October 2021 <https://stop-it-project.eu/download/continuity-plan-for-the-training-courses-and-services-d8-4/#>
- [6] Dionysios Nikolopoulos, Georgios Moraitis, and Christos Makropoulos: "Chapter 7, Strategic and Tactical Cyber-Physical Security for Critical Water Infrastructures", in Soldatos, Praça, Jovanović (eds.) (2021), "Cyber-Physical Threat Intelligence for Critical Infrastructures Security: Securing Critical Infrastructures in Air Transport, Water, Gas, Healthcare, Finance and Industry", Boston-Delft: now publishers, <http://dx.doi.org/10.1561/9781680838237>].
- [7] Nikolopoulos, D.; Moraitis, G.; Bouziotas, D.; Lykou, A.; Karavokiros, G.; Makropoulos, C. Cyber-Physical Stress-Testing Platform for Water Distribution Networks. *J. Environ. Eng.* 2020, 146, 04020061, doi: 10.1061/(ASCE)EE.1943-7870.0001722
- [8] John Soldatos (ed.), Isabel Praça (ed.), Aleksandar Jovanović (ed.) (2021), "Cyber-Physical Threat Intelligence for Critical Infrastructures Security: Securing Critical Infrastructures in Air Transport, Water, Gas, Healthcare, Finance and Industry", Boston-Delft: now publishers, <http://dx.doi.org/10.1561/9781680838237>
- [9] Gill, H. A Continuing Vision: Cyber-physical Systems. In Proceedings of the HCSS National Workshop on New Research Directions for High Confidence Transportation CPS: Automotive, Aviation, and Rail; Washington, DC, USA, 2008; pp. 1–28.
- [10] Nikolopoulos, D.; Makropoulos, C.; Kalogeras, D.; Monokrousou, K.; Tsoukalas, I. Developing a Stress-Testing Platform for Cyber-Physical Water Infrastructure. In Proceedings of the 2018 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater); IEEE, 2018; pp. 9–11
- [11] Irmak, Erdal, and İsmail Erkek. "An overview of cyber-attack vectors on SCADA systems." 2018 6th international symposium on digital forensic and security (ISDFS). IEEE, 2018 <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8355379&tag=1>

- [12] Riccardo Taormina; Stefano Galelli, Nils Ole Tippenhauer; Elad Salomons; and Avi Ostfeld, "Characterizing Cyber-Physical Attacks on Water Distribution Systems", *Journal of Water Resources Planning and Management*, Volume 143, Issue 5 - May 2017 <https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29WR.1943-5452.0000749>
- [13] Amin, Saurabh, Xavier Litrico, Shankar Sastry, and Alexandre M. Bayen "Cyber security of water SCADA systems—Part I: Analysis and experimentation of stealthy deception attacks." *IEEE Transactions on Control Systems Technology* 21.5 (2012): 1963-1970. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6303885>
- [14] Amin, S., Litrico, X., Sastry, S. S., & Bayen, A. M. Stealthy deception attacks on water SCADA systems. In *Proceedings of the 13th ACM international conference on Hybrid systems: computation and control* (pp. 161-170). (2010, April) <https://dl.acm.org/doi/pdf/10.1145/1755952.1755976>
- [15] Klise, K.A.; Bynum, M.; Moriarty, D.; Murray, R. A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. *Environ. Model. Softw.* 2017, 95, 420–431, doi: 10.1016/j.envsoft.2017.06.022
- [16] Yulia Cherdantseva, Pete Burnap, Andrew Blyth, Peter Eden, Kevin Jones, Hugh Soulsby, Kristan Stoddart: A review of cyber security risk assessment methods for SCADA systems, *Computers & Security*, Volume 56, 2016, Pages 1-27, <https://doi.org/10.1016/j.cose.2015.09.009>.
- [17] Robles, Roslin John, and Min-kyu Choi. "Assessment of the vulnerabilities of SCADA, control systems and critical infrastructure systems." *International Journal of Grid and Distributed Computing* 2.2 (2009): 27-34.
- [18] Eric Luijff, SCADA Security Good Practices for the Drinking Water Sector, TNO Report TNO-DV 2008 C096
- [19] Alquwatli, Mohammed H., Mohamed Hadi Habaebi, and Sheraz Khan. "Review of SCADA Systems and IoT Honey pots." 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS). IEEE, 2019. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9117330>
- [20] Zhu, Bonnie, Anthony Joseph, and Shankar Sastry. "A taxonomy of cyber attacks on SCADA systems." 2011 International conference on internet of things and 4th international conference on cyber, physical and social computing. IEEE, 2011. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6142258>
- [21] Noam Erez, Avishai Wool: Control variable classification, modeling and anomaly detection in Modbus/TCP SCADA systems, *International Journal of Critical Infrastructure Protection*, Volume 10, 2015, Pages 59-70, <https://doi.org/10.1016/j.ijcip.2015.05.001>.
- [22] R. Negi, P. Kumar, S. Ghosh, S. K. Shukla and A. Gahlot, "Vulnerability Assessment and Mitigation for Industrial Critical Infrastructures with Cyber Physical Test Bed," *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, 2019, pp. 145-152, <https://doi.org/10.1109/ICPHYS.2019.8780291>
- [23] A. O. Gomez Rivera and D. K. Tosh, "Towards Security and Privacy of SCADA Systems through Decentralized Architecture," 2019 International Conference on Computational Science and Computational Intelligence (CSCI), 2019, pp. 1224-1229, <https://doi.org/10.1109/CSCI49370.2019.00230>
- [24] I. Darwish, O. Igbe, O. Celebi, T. Saadawi and J. Soryal, "Smart Grid DNP3 Vulnerability Analysis and Experimentation," 2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing, 2015, pp. 141-147, <https://doi.org/10.1109/CSCloud.2015.86>

- [25] S. Adepur, J. Prakash and A. Mathur, "WaterJam: An Experimental Case Study of Jamming Attacks on a Water Treatment System," 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), 2017, pp. 341-347, <https://doi.org/10.1109/QRS-C.2017.64>
- [26] Jaatun, Martin Gilje; Wille, Egil; Bernsmed, Karin; Kilskar, Stine Skaufel: " Grunnprinsipper for IKT-sikkerhet i industrielle IKT-systemer", SINTEF Report 2021:00055 (in Norwegian) <https://sintef.brage.unit.no/sintef-xmlui/handle/11250/2835081>
- [27] Ugarelli, Rita; Bosco, Camillo; Jaatun, Martin and Bour, Guillaume: DWC D4.3: Security assessment of cyber-physical flow of information in strategic, tactical and operational dimensions regarding DWC digital solutions, 2022
- [28] M. Ahmadi et al., D4.4: Cyber-Physical threats stress-testing platform, STOP-IT (2019).
- [29] "What You Need To Know About the SolarWinds Supply-Chain Attack | SANS Institute." <https://www.sans.org/blog/what-you-need-to-know-about-the-solarwinds-supply-chain-attack/> (accessed May 02, 2022).
- [30] "European Banking Authority hit by Microsoft Exchange hack - BBC News." <https://www.bbc.com/news/technology-56321567> (accessed May 02, 2022).
- [31] "Norway's parliament hit by new hack attack," Reuters, Mar. 10, 2021. Accessed: May 02, 2022. [Online]. Available: <https://www.reuters.com/article/us-norway-cyber-idUSKBN2B21TX>
- [32] "A Large-Scale Supply Chain Attack Distributed Over 800 Malicious NPM Packages," The Hacker News. <https://thehackernews.com/2022/03/a-threat-actor-dubbed-red-lili-has-been.html> (accessed May 02, 2022).
- [33] "Governments need to reassess security infrastructures | Orange Business Services." <http://www.orange-business.com/en/magazine/new-generation-critical-infrastructures-secure> (accessed May 02, 2022).
- [34] "Clear the 'air gap' myth to evade cyber threats - Securing critical infrastructure in the digital world," Nokia. <https://www.nokia.com/networks/insights/critical-infrastructure-enterprise-security/> (accessed May 02, 2022).
- [35] S. H. Weingart, "Physical Security Devices for Computer Subsystems: A Survey of Attacks and Defences," in Proceedings of the Second International Workshop on Cryptographic Hardware and Embedded Systems, Berlin, Heidelberg, Aug. 2000, pp. 302–317.
- [36] Microsoft, "Ten Immutable Laws Of Security (Version 2.0)." <https://docs.microsoft.com/en-us/archive/blogs/rhalebheer/ten-immutable-laws-of-security-version-2-0> (accessed Apr. 03, 2022).
- [37] 'Vulnerabilities and Exploits', ENISA. <https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/vulnerabilities-and-exploits> (accessed Apr. 03, 2022).
- [38] B. Schneier, 'Essays: The Process of Security - Schneier on Security'. https://www.schneier.com/essays/archives/2000/04/the_process_of_secur.html (accessed Apr. 03, 2022).
- [39] C. Johnson, M. Badger, D. Waltermire, J. Snyder, and C. Skorupka, 'Guide to Cyber Threat Information Sharing', National Institute of Standards and Technology, NIST Special Publication (SP) 800-150, Oct. 2016. doi: 10.6028/NIST.SP.800-150.
- [40] 'Ransomware Attack Topples Telemarketing Firm, Leaving Hundreds Jobless | Threatpost'. <https://threatpost.com/ransomware-attack-topples-telemarketing-firm/151530/> (accessed May 29, 2022).
- [41] NIST CSRC, 'Principle of Least Privilege'. https://csrc.nist.gov/glossary/term/principle_of_least_privilege (accessed Apr. 04, 2022).

- [42] 'Company shuts down because of ransomware, leaves 300 without jobs just before holidays | ZDNet'. <https://www.zdnet.com/article/company-shuts-down-because-of-ransomware-leaves-300-without-jobs-just-before-holidays/> (accessed May 29, 2022).
- [43] 'TV station gets hacked. And then broadcasts its passwords in a report about the hack', The Independent, Apr. 12, 2015. <https://www.independent.co.uk/tech/tv5monde-hack-staff-accidentally-show-passwords-in-report-about-huge-cyberattack-10168475.html> (accessed May 29, 2022).
- [44] H. Adkins, B. Beyer, P. Blankinship, A. Oprea, P. Lewandowski, and A. Stubblefield, Building Secure and Reliable Systems: Best Practices for Designing, Implementing, and Maintaining Systems. O'Reilly Media, 2020. [Online]. Available: <https://books.google.no/books?id=Kn7UxwEACAAJ>
- [45] 'Baseline Security Recommendations for IoT', ENISA. <https://www.enisa.europa.eu/publications/baseline-security-recommendations-for-iot> (accessed May 29, 2022).
- [46] H. Schwarzmüller, A. Vennesland, P. H. Haro, and G. Bour, "D4.1: Interoperable and Secure Flow of Information - Cyber-physical Sphere and Interoperability Aspects in the Utilities Regarding the DWC Solutions," Mar. 2021, doi: 10.5281/zenodo.6497313.
- [47] Haro, P.H., Hanssen, B.J. and Vennesland, A. "D4.7 Semantic interoperable middleware - final version" DWC, 2020.
- [48] Ekaputra, Fajar, et al. "Ontology-based data integration in multi-disciplinary engineering environments: A review." Open Journal of Information Systems 4.1 (2017): 1-26.
- [49] Chen, J., Hu, P., Jimenez-Ruiz, E., Holter, O. M., Antonyrajah, D., & Horrocks, I. (2021). Owl2vec: Embedding of OWL ontologies. Machine Learning, 110(7), 1813-1845.
- [50] Vennesland, A., Haro, P.H. and Hanssen, B.J. "D4.4: Semantic Interoperability Design Requirements" 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.4320417>.
- [51] Pijuan, J. et al. "B-WaterSmart, FIWARE based interoperability framework – deliverable 3.1" 2022.
- [52] Slay, Jill, and Michael Miller. "Lessons learned from the Maroochy water breach." *International conference on critical infrastructure protection*. Springer, Boston, MA, 2007. https://link.springer.com/chapter/10.1007/978-0-387-75462-8_6
- [53] Liebowitz, Matt "Hacker says he breached Texas water plant network", NBC News, Nov 22, 2011, <https://www.nbcnews.com/id/wbna45394132>
- [54] Vera, Amir; Jamiel Lynch and Christina Carrega, "Someone tried to poison a Florida city by hacking into the water treatment system, sheriff says", CNN, February 9, 2021 <https://edition.cnn.com/2021/02/08/us/oldsmar-florida-hack-water-poison>
- [55] Dragos Inc.: "CRASHOVERRIDE - Analysis of the Threat to Electric Grid Operations", <https://www.dragos.com/wp-content/uploads/CrashOverride-01.pdf>
- [56] Vennesland, Audun: DWC D4.5: DWC Water Value Chains Ontology (v1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.6497063>, 2021
- [57] Bour, Guillaume; Selseth, Ingrid; Jaatun, Martin and Ugarelli, Rita: DWC D4.2: Risk Identification Database & Risk Reduction Measures Database, <https://zenodo.org/record/6497050>, 2021
- [58] Kerckhoffs, Auguste. La cryptographie militaire, *Journal des sciences militaires*, Vol. IX, p. 5-38, janvier, 1883. <http://www.bibnum.education.fr/calculinformatique/cryptologie/la-cryptographie-militaire>
- [59] International Organization for Standardization: "Risk Management – Guidelines", ISO 31000:2018 <https://www.iso.org/standard/65694.html>

- [60] Salomons, Elad; Vila, Aleix; Caubet, Juan; Meseguer, Jordi; Cembrano, Gabriela; Bonet, Enric; Diaz, Rodrigo; Gonzalez, Gustavo; Bernsmed, Karin: STOP-IT D5.4 Toolbox of technologies for securing IT and SCADA systems in CI - Supporting document <https://stop-it-project.eu/results/toolbox-for-it-and-scada-security/>
- [61] Ugarelli, R., Koti, J., Bonet, E., Makropoulos, C., Caubet, J., Camarinopoulos, S., ... & Jaatun, M. G. (2019). STOP-IT-Strategic, Tactical, Operational Protection of water Infrastructure against cyber-physical Threats. Physical and Cyber Safety in Critical Water Infrastructure, 56, 130.
- [62] Stein, Ulf et al.: Digitalisation in the water sector - Recommendations for policy developments at EU Level, European Research Executive Agency (European Commission), October 2022, <https://op.europa.eu/en/publication-detail/-/publication/6eb837b2-54df-11ed-92ed-01aa75ed71a1>
- [63] Selseth, Ingrid and Ugarelli, Rita: InfraRisk-CP - User's guide, STOP-IT Project, June 2020 <https://zenodo.org/record/7347820#.Y3zj8X3MLuo>
- [64] Cybersecurity and Infrastructure Security Agency (CISA): Cyber Risks Resources for the Water and Wastewater Systems Sector, Infographic, October 2021, <https://www.cisa.gov/sites/default/files/publications/infographic-manage-wastewater-national-critical-function-102021-508.pdf>

Appendix A: Sample chapters

In the following we present sample chapters for the training material, some of them slightly modified text taken from the STOP-IT project [5].

A.1. Sample material for chapter: Introduction

Topic

This course module serves as an introduction to the DWC training material/course.

Goals

- To motivate why security of Cyber-Physical Systems is an important area in the water domain
- To provide examples of past incidents
- To provide an outline of the training material

Text

One of the earliest cases of cyber-physical security breach was the Maroochy Shire incident [52], where disgruntled former employee Vitek Boden abused credentials for wireless SCADA controllers, causing vast amounts of untreated sewage to be dumped in the environment, resulting in damages in the range of AUD 1 million. You could argue that this was more a case of poor configuration control than a cyber vulnerability (since Boden's credentials and access to equipment should have been revoked), but history does not stop here.

About a decade later, an image purporting to be a screenshot of a water treatment plant HMI was published on the text-sharing site Pastebin [53]. For people without local knowledge, it might have been a little confusing, since the image was headed "The city of South Houston Nevada water plant" – and Houston is surely in Texas, not in Nevada? A search with an online map tool quickly shows that the city of South Houston is indeed a suburb of the more famous city of Houston, a short 20-minute car ride away. Further searches reveal that the city of South Houston has a *street* named Nevada, where at the intersection of Beaumont St there appears to be a vacant lot. Switching to aerial photography it was possible to see something resembling a water tank, which was confirmed when switching to street view.

The technology enthusiast in question, simply referred to as "Pr0f", claimed to have accessed the HMI of several Siemens Simatic PLCs connected to the internet, due to the default password of "100" not having been changed. Ignoring for a moment the ridiculously short and simple default password, relying on an attacker not knowing the default password is the very worst form of security by obscurity. Used Siemens PLCs are abundantly available for purchase via online marketplaces, and even without buying your own PLC, default passwords tend to be shared widely in online forums.

Putting an industrial control system on the internet can seldom be done with impunity; there is even a specialized search engine that can be used to find any Internet-of-Things type of device connected to the internet (e.g., web cameras, temperature sensors, Programmable Logic Controllers). In February 2021, the threat to water infrastructures became real when an unauthorized person gained remote access to the water treatment plant in the town of Oldsmar, Florida, and managed to increase the amount of lye added to the drinking water to a potentially dangerous level [54]. Personnel at the water treatment plant detected the change immediately and was able to reset the system to a safe state. In this case, the attack was made possible due to a remote desktop application (TeamViewer) being accessible from the open internet.

Other industries

With a sufficiently motivated and technologically capable adversary, hardly any critical infrastructure is impervious to attack. This was evident in the attacks against the Ukrainian power grid in 2015 and 2016 [55], where many people were left in the dark for several hours and more, and where restoration in many cases was only possible through manual intervention.

Fundamental problems

In computer security, "security through obscurity" is something that is considered a poor substitute for real security. The Dutch cryptographer Auguste Kerckhoffs formulated what has become known as "Kerckhoffs' principle", which states that the security of a cryptographic device should only rely on the secrecy of the key, not on keeping the mechanism or device secret [58]. As Kerckhoffs' contemporaries in the 19th century no doubt discovered, cryptographic devices tend to get captured by the enemy, and it quickly becomes tedious having to replace all equipment each time this happens.

Kerckhoffs' principle is equally applicable to process control systems in the water sector, so if your security relies on the bad guys not getting their hands on your particular brand of PLC, you are doing it wrong. Also note that security generally also relies on proper configuration; default usernames and passwords are frequently documented in user manuals, and these are typically available for download on the internet, as was the case in South Houston (see above).

Structure of the training material

The remainder of the training material is structured as follows:

- Water systems as cyber-physical entities
Provides an introduction to how and why utilities are going digital, in what ways (sensors, etc.) and how that means more CP risk, and offers a basic overview of how a water utility functions as a cyber-physical system (two interconnected layers, sensor and actuator principle). Further examples of CPS attacks are provided: cyber-attacks, physical impacts (and vice versa).
- Information Systems for water

Presents key information systems for water, with a special focus on SCADA units. Covers attack vectors relevant to these information systems (including physical, cyber and cyber-physical attacks).

- Communication technologies for water

Presents key communication technologies for (smart) water, wired and/or wireless, and the specific attack vectors relevant to these communication technologies (as above).

- Risk Management for Cyber-physical Security

Introduces the risk management process described by the ISO 31000-2018 standard and shows how it relates to cyber-physical risk assessment management. Introduces a cyber-physical risk assessment tool for the water sector.

- IoT Security for water systems

Introduces the concept of IoT security for water systems and presents a concrete methodology for security testing of IoT devices, providing some "dos and don'ts" and good practice in form of a checklist.

- Organizational Resilience Training

Introduces the online Training for Operational Resilience Capabilities (TORC) game, and documents how to play TORC as a teaching experience.

A.2. Sample material for chapter: Water systems as cyber-physical entities

Topic

The topic covered in this course module is on treating water systems as cyber-physical systems (CPS), with an introduction of the main risks associated with CPS.

Note: This module is based on material from STOP-IT [5], inspired from contents in relevant NTNU courses and a book chapter on CPS [6].

Goals

- To explain, in very simple terms, why and how water systems can be considered CPS and further explain their layers (cyber and physical), as well as their interaction.
- To explain the **risks** associated with these systems. Emphasize the links with STOP-IT tools.

Instructor pre-requisites

This topic should be taught by an instructor with competence in water network engineering and digital tools connected to this domain.

Text

Water systems generally include physical and cyber elements as part of their network. Physical elements include all assets of water collection, treatment and distribution needed to bring water in a safe and reliable way to customers, such as dams or reservoirs, spillways, water treatment plant buildings and technologies, aqueducts and tanks and, eventually, pipes that deliver water to end users. Cyber elements include, but are not limited to, all types of water quantity and quality sensors, the Supervisory Control and Data Acquisition (SCADA) system that monitors and controls processes, and networking elements to connect these elements (lines, LAN/WAN networks etc.). Both layers (i.e., the physical and the cyber) can be studied and designed separately in terms of functionality and risks; however, it is also important to consider these two layers as one interconnected (i.e., cyber-physical) entity for water, where cyber elements affect their physical counterparts and vice versa.

A system that integrates physical processes with computational engineering systems is termed a cyber-physical system (CPS). The cyber layer of this integration employs a networking, computing, and communication core of embedded computers and devices that monitors, controls and coordinates the physical processes. This synergy is accomplished via feedback loops, where the outcome of a physical process affects computation and vice versa [8]. While the term CPS was introduced in 2008 to describe “deeply embedded” systems that are fully integrated hybridizations of computational (logical) and physical actions [9], the concept and its application has started long before that, with the onset of automated control systems for physical processes and the handling of digital information by

mainframe computers. Contemporary CPSs are evolving, rapidly benefiting from the emergence of other related technologies in the informatics and computer science fields, such as IoT (internet of things), big data, cloud computing, novel sensor technology, and other advances in ICT like optical fiber wire connections and 5G cellular connectivity. Essentially all smart water systems can be considered CPS, as they rely heavily on the cyber layer and its interaction with the physical one.

A basic principle to understand the interaction between the cyber and physical layers in a CPS is the 'sensor-actuator' principle. According to this principle, what is commonly found in water systems is that cyber elements (sensors) gather information, in (near) real-time, about aspects of the water system. Operators then use this information, either by decision-making or through automation, to decide about the status and operation of physical elements, such as valves, gates, spillways etc. One can only then extrapolate that any compromise to a cyber element, such as a sensor, will have direct, physical impacts to the water system, as it will affect the operation of its assets.

Of course, a water CPS is not only contained to sensors and actuators. A plethora of basic elements for a CPS exist, such as:

- Conversion Units, such as Remote Terminal Units (RTU) and Programmable Logic Controllers (PLCs). These are connected to sensors and interpret the data collected, convert it to digital (if the sensor is analogue) and can command actuators using logical rules with the data collected. These units also send the data to the central SCADA unit.
- Master Units, the most important of which is the Supervisory Control and Data Acquisition (SCADA) unit. This is essentially the most prominent part of the cyber layer and the backbone of every CPS, as it serves as the central monitoring and control unit, gathering data from all distributed sub-units (RTUs/PLCs) and presenting an overview of the system status to the operators. The SCADA generally includes a database of all data collected (Historian), as well as a Human Machine Interface (HMI).
- Communication Networks and Protocols, which are the hardware that connects information across peripherals in the system (e.g., between a sensor, a PLC and an actuator), but also from distributed elements to the central SCADA unit. Communication networks include both wired (telephone lines, WAN circuits, fiber-optic cables) and wireless technologies (Wi-Fi, Bluetooth, radio, cellular, satellite), as well as the required protocols for device interaction (e.g., TCP/IP, Modbus).

An overview of these elements can be seen in Figure 2, where one may see a water distribution network with a single source, comprising pumps and valves in a city (lower part). There are sensors at key parts of the network, measuring operational attributes such as the water pressure and flow at given intervals. This information is then passed to RTUs, which act as peripheral information collection terminals that then send this information (through a wireless or wired manner) to the main SCADA unit. The SCADA unit forms the center of operations, including all monitoring, analysis and customer service components. Finally, operator decisions travel the inverse way: from the SCADA unit back to the relevant PLCs that act as the convertors of digital information to physical actions through relevant switches. The result is a physical action in an actuator, for instance a change in pump settings or the closure of a valve. This flow of information towards the SCADA unit and back happens in (near) real-time, so a wealth of information is acquired during normal water system operations, aggregated within the SCADA unit and used to detect anomalies, assist operator decisions, improve customer service etc.

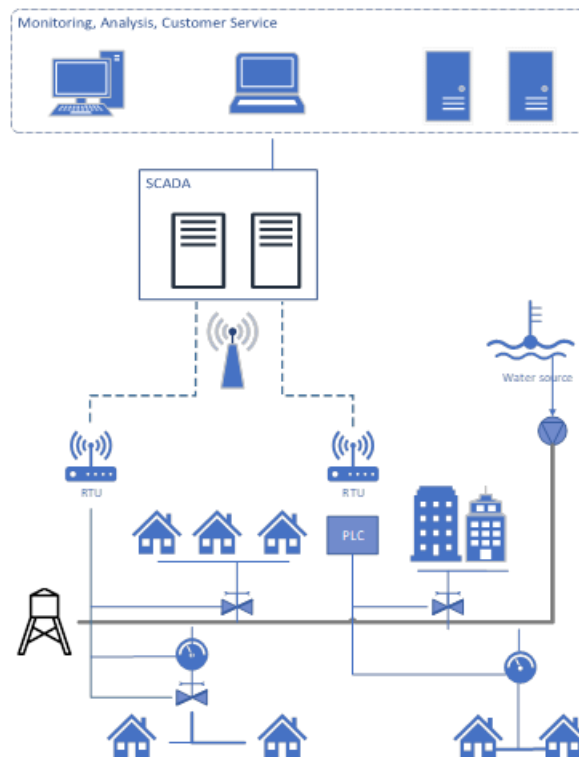


Figure 2: Overview of the interactions of cyber and physical elements in a CPS

The various aforementioned elements in a CPS allow multiple possible attack routes of cyberattacks for adversaries. Physical attacks are the simplest form of CPS attacks, such as tampering with field devices or modifying and compromising wired connections (e.g., cable damage). Another common cyber attack vector is the exploitation of backdoor (i.e., unauthorized hidden software or hardware mechanisms to circumvent security measures) or unintentional security holes in the network perimeter that allows some form of remote access or control. Further, attacks within the SCADA system may occur, for instance targeting the CPS database with methods like SQL injection, where malicious code is inserted in queries to manipulate data or even controls of the system. Finally, wireless networks have their own cybersecurity concerns, including eavesdropping on information from unsecured networks, compromise of remotely controlled, wireless sensors and actuators, and jamming signals. The communication hijacking between such components (from a signal source to a destination) constitutes a wide class of attacks, called Man-in-the-middle, where the attacker may try to (a) interrupt a message so that data are not received at the destination, (b) intercept a message for information eavesdropping, (c) modify the data of a message, so that an altered version is received at the destination, or (d) by imposing the source, fabricate a bogus (fake) message, and send it to the destination. As this information is used for actions on the physical assets of the water system (e.g., through PLCs), this form of attacks is particularly severe for water systems. Cyberattacks on CPSs can be potentially even more hazardous when coupled with physical attacks (sabotage or other deliberate malicious actions) in a combined cyber-physical attack. For example, in a water CPS, adversaries may perform a terrorist attack such as contaminating a water source and simultaneously perform a cyberattack that manipulates input data from water quality sensors to magnify impact.

To summarize, attacks to water systems as CPS can have the following threats' nature:

- Cyber: Voluntary or not intent of individuals or groups to electronically corrupt or seize control of data or information essential to system operations.
- Physical: Water infrastructure is prone to any kind of physical threats either due to natural hazards (earthquakes, floods, etc.) or terrorist attacks or even due to an accident. The threat is a physical occurrence on the water supply system. By the physical type of threats, assets or technical devices of the water supply system will be damaged or manipulated. The physical threat may also destroy or damage sensors, data transmission lines or the process control/SCADA system in a way that the normal function is no longer possible.
- Cyber-Physical: The threat has a combined cyber-physical nature. It can be generated in different ways, such as:
 - Combined cyber-physical threats: coordinated and long-term attacks to the critical infrastructure (CI) to reach and compromise the normal functioning.
 - Cyber threat to any of the physical components of the water infrastructure, e.g. monitoring devices (including e.g. IP cameras, networked sensors, Automatic Meter Reading (AMR)/Advanced Metering Infrastructure (AMI)) that become more vulnerable to cyber-attacks due to their higher automation/networking level
 - Physical threats to the “cyber” environment of the water utilities, e.g. Intrusion of attackers to the utilities control & operation centers (access to computers) or SCADA devices, etc.

EXAMPLE of relevant cyber attack

Cyber attacks are becoming more and more frequent. There are examples of events in which attackers successfully deployed ransomware within a water utility’s Supervisory Control and Data Acquisition (SCADA) system, forcing the facilities to switch to manual operation. Ransomware is most commonly deployed against information technology (IT) and business operations systems, but ransomware can also “infect connected Operational Technology (OT) systems, particularly if there is not adequate segmentation between IT and OT systems” [64]. Information about events is restricted, except for some very well known ones, as described in the following.

In February 2021, the City of Oldsmar, Florida, suffered an attack [54] that could have compromised public health. A hacker breached the network of the city’s drinking water treatment facility and manipulated the levels of chemicals used in the water purification process, attempting to increase the concentration of sodium hydroxide from its normal 100 parts-per-million (ppm) to 11,100 ppm. Fortunately, an employee detected the hacker’s movements in real time and stopped the chemicals from being released into the water supply.

The officials noted that it would have taken 24 to 36 hours for the chemicals to contaminate the water supply. The officials acknowledged, however, that the employee who witnessed the intrusion initially failed to report it, assuming it was another employee remotely accessing the network through an older program, rather than a hacker. The FBI cited poor cybersecurity, including weak passwords and outdated operating systems, as contributors to the hacker’s success

A similar attack succeeded in 2019 in shutting down the treatment processes at a drinking water plant in Kansas. The Department of Justice accused a former employee of intentionally threatening public health and safety. Despite having resigned from the company two months earlier, the employee used his still-active remote-access credentials to interfere with the system.

Modeling approaches

Perceiving a water system as a CPS is no easy task, and needs the support of tools and models that help explore the effect of cyber-physical attacks on systems (and the cascade of effects between the two layers). Recent research has produced a variety of cyber-physical tools, which can be classified into two categories with regards to the representation of the cyber layer: (i) emulation/virtualization based and (ii) simulation based. The first category (emulation/virtualization) formulates a detailed model of the cyber layer of the water CPS. This provides high fidelity in the explicit modeling of the behavior of any real or virtual cyber component (from network cables to software protocols), using emulator platforms, discrete event simulators, virtualization machines, and software defined networks (SDNs). However, the implemented models tend to be domain-specific and applicable only to a specific CPS, with almost no chance of scalability or transferability to other systems. Moreover, monetary and time budget constraints increase with the scale of the systems and may be prohibitive for smaller utilities [10]. The second approach (simulation) represents both the cyber and physical layers with simulation models. As such, programming functions, routines, classes, and data structures represent elements and functionality of the cyber layer, modeling the information flow with feedback loops and interactions between the cyber and physical layers. This results in a lower fidelity process, since the focus is on the outcome of a cyber-operation or the state of a cyber-component, without the need for “bit-wise” modeling of interaction. Advantages compared to emulation/virtualization approaches include (a) “what-if” scenarios of cyber-physical attacks can be assessed without limitations, from the perspective of the water utility and by risk management officers untrained in ICT/IT fields and (b) the coupling with physical process simulators/models is much easier via the use of software wrappers, application programming interfaces, or dynamic link libraries. Simulation-based tools are also more generic and thus applicable to multiple water utilities, as long as their system can be converted to a network topology (with a physical and cyber layer) used by the model.

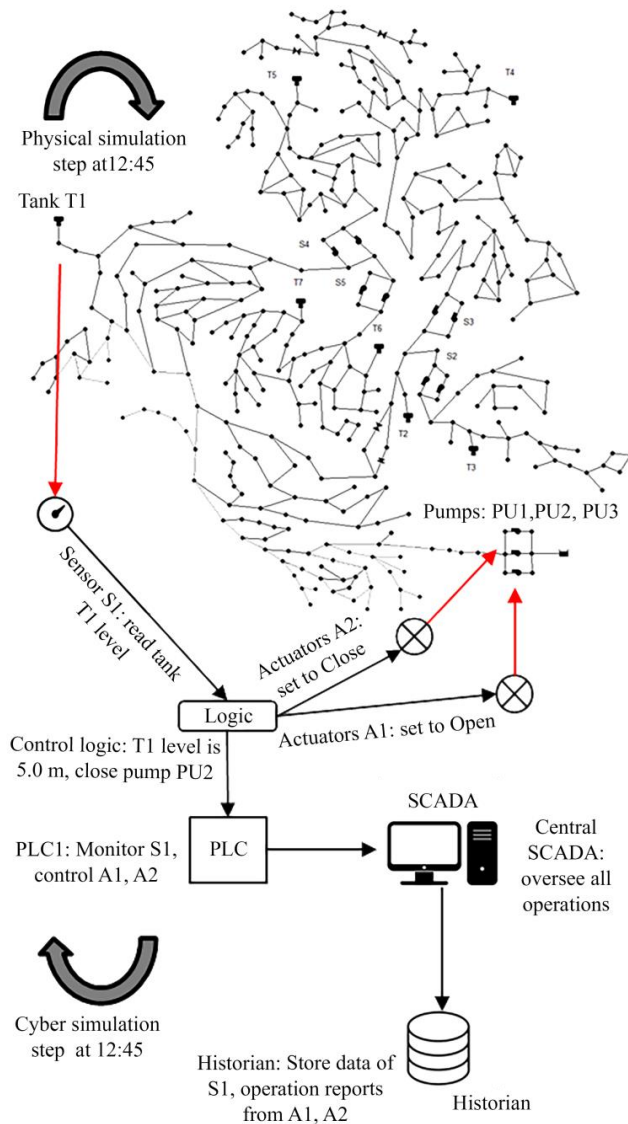


Figure 3: Example of the RISKNOUGHT simulation-based tool for CPS and the different elements modeled with it.

A simulation-based tool that is developed within STOP-IT is RISKNOUGHT (Figure 3), a holistic cyber-physical stress testing platform developed in Python [7]. The platform represents any water distribution system as a CPS, via automatically formulating a customizable SCADA model with enhanced control logic (e.g., users can add controls for water quality contamination response measures, controls based on data from the operational historian, etc.). An attack module is used to devise scenarios of complex cyber-physical attacks, for example, combinations of cyberattacks and backflow contaminant injection attacks. The latest version of RISKNOUGHT is interfaced with the water distribution model EPANET 2.2 to simulate the physical layer, and leverages the WNTR water network resilience analysis Python package as a python interface [15].

Evaluation questions

1.) A contemporary water system includes:

<input type="checkbox"/>	A: Physical elements, such as pipes and valves.
<input type="checkbox"/>	B: Cyber elements, such as SCADA units and digital sensors.
<input checked="" type="checkbox"/>	C: Both physical and cyber elements.
<input type="checkbox"/>	D: Neither physical nor cyber elements.

Note: C is correct

2.) The concept of Cyber-Physical Systems (CPS) in water means that one has to look at a water system as a:

<input type="checkbox"/>	A: Set of two connected layers, physical and cyber, where the cyber layer interacts with the physical layer.
<input checked="" type="checkbox"/>	B: Set of two connected layers, physical and cyber, where both layers interact and exchange information with each other.
<input type="checkbox"/>	C: Set of two connected layers, physical and cyber, where the physical layer interacts with the cyber layer.
<input type="checkbox"/>	D: Set of multiple connected layers, including physical and digital ones but also human capital and financial ones.

Note: B is correct

3.) The most important cyber element in a water CPS is:

<input type="checkbox"/>	A: A sensor that measures a critical aspect of the network operation (e.g. flow in a District Meter Area).
<input type="checkbox"/>	B: A PLC, because it converts digital information, through logic rules, to physical actions.
<input checked="" type="checkbox"/>	C: The SCADA unit, as it aggregates information from all sources and is a fundamental tool of the system operators.
<input type="checkbox"/>	D: The Historian database because it includes a thorough history of network operations and data.

Note: C is correct

4.) The sensor-actuator principle in CPS means that:

<input checked="" type="checkbox"/>	A: Information from cyber elements (from sensors) affects physical actions (through actuators).
<input type="checkbox"/>	B: Information from physical elements (from actuators) affects cyber actions (sensors).
<input type="checkbox"/>	C: There are both sensors and actuators in all water systems everywhere.
<input type="checkbox"/>	D: There are either sensors or actuators in all water systems everywhere.

Note: A is correct

5.) A simulation-based CPS tool is:

<input type="checkbox"/>	A: High-fidelity, as there is explicit modeling of the (bit-wise) behavior of any real or virtual cyber component (from network cables to software protocols).
<input checked="" type="checkbox"/>	B: Low-fidelity, as the focus is on the overall outcome of a cyber-operation or the state of a cyber-component, without the need for lower-level, “bit-wise” modeling of element interactions.
<input type="checkbox"/>	C: Either high-fidelity or low-fidelity, depending on the simulation used.
<input type="checkbox"/>	D: Non-fidelity, as it does not combine the physical and cyber layers.

Note: B is correct

6.) An example of an attack in a CPS is:

<input type="checkbox"/>	A: An eavesdropping attack, compromising information between a sensor and a PLC.
<input type="checkbox"/>	B: An SQL injection attack in the SCADA database
<input type="checkbox"/>	C: A physical attack to a wired connection within the water system.
<input type="checkbox"/>	D: A malware installed in the SCADA unit using a security loophole.
<input type="checkbox"/>	E: Answers (A), (B) and (D).
<input checked="" type="checkbox"/>	F: Answers (A), (B), (C) and (D).
<input type="checkbox"/>	G: Answers (A) and (D).

Note: F is correct

7.) An example of an attack in a CPS according to the sensor-actuator principle is:

<input type="checkbox"/>	A: An SQL injection attack in the SCADA database.
<input type="checkbox"/>	B: A physical attack damaging a fiber optic cable connection within the water system.
<input type="checkbox"/>	C: A malicious change in the logic rules of a PLC, so that a valve of a network does not close at regular sensor reading thresholds.
<input type="checkbox"/>	D: A compromise of the communication system between the sensor and the PLC, so that bogus (fake, synthetic) signals are read by the PLC instead of real information coming from the sensor.
<input type="checkbox"/>	E: Answers (A), (C) and (D).
<input checked="" type="checkbox"/>	F: Answers © and (D).
<input type="checkbox"/>	G: Answers (A), (B), (C) and (D).

Note: F is correct

8.) What are the main categories of CPS modeling tools and their main characteristics?

.....

Note: emulation-based vs. simulation-based, mention of fidelity aspects

9.) Describe two basic elements in a water CPS and mention if they are (primarily) cyber or physical.

.....

Note: Based on the text mention, including extra elements (SCADA, communication protocols etc.)

A.3. Sample material for chapter: Information systems for water

Topic

The topic covered in this course module is on **information systems** in (smart) water systems, as well as their main risks.

Note: This module is based on material from STOP-IT [5].

Goals

- Explain, in very simple terms, key information systems associated with (smart) water systems (wired and wireless), e.g., SCADA systems etc.
- Explain the **risks** associated with these systems. Emphasize the links with STOP-IT tools (e.g., mention jamming attacks).

Instructor pre-requisites

This topic should be taught by an instructor with competence in computer network engineering.

Text

Introduction

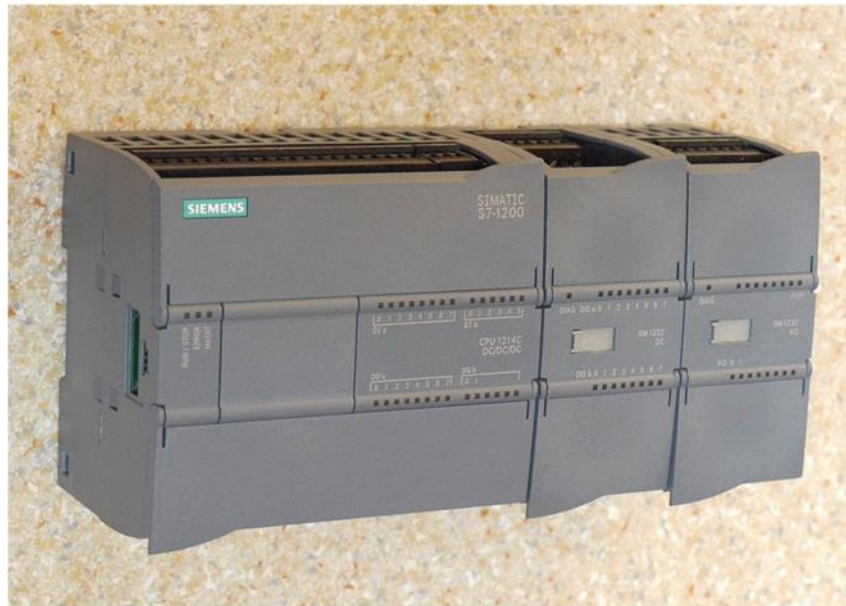
In this module we will describe some of the information systems commonly used in (smart) water systems and some major problems associated with them. We will also give a link to some of the tools available in the STOP-IT toolbox.

Water systems are distributed systems, covering a wide geographical area. To monitor and control the processes of this critical infrastructure Supervisory Control and Data Acquisition (SCADA) systems are used. SCADA systems consist of multiple devices such as field devices, local processors, and Human Machine Interfaces (HMIs) that allow for real time monitoring and controlling of the water system. The implementation of the SCADA systems varies from water system to water system. Figure 5 illustrates a typical SCADA system architecture.

Supervisory Control and Data Acquisition

Field devices such as sensors, actuators, or control devices, are used to measure and collect signals from different parts of the water system. Further, these signals are sent to a local processor such as a Remote Terminal Unit (RTU) or a Programmable Logic Controller (PLC). The RTU scans and collects the incoming analogue signals and converts them to digital data which is sent to the control center. The RTU also receives digital commands from the control center and handles alarms. PLCs monitor sensors and use the data to control valves and other actuators. The decision made by the PLC is based on the

sensor data along with a user created program. Figure 4 shows an example of a Simatic S7-300 PLC from Siemens. The control center is the core of the SCADA system. It includes a Master Terminal Unit (MTU), which issues commands to and gathers data from RTUs. The MTU also stores process data in order to display information to human operators to support decision making. The information is displayed on HMIs, where operators can interact with the supervisory system to monitor and control the processes of the system [16]. The control center allows the operators to receive and assess alarms, analyse data, and send control signals to actuators. The central computer also allows for storage of measurement data and system conditions in a database and for the generation of reports.



"File:Simatic S7-1200.JPG" by UlrichAAB is marked with CC BY-SA 3.0.

Figure 4: PLC from Siemens

If the control center is not protected by firewalls, security patches or intrusion prevention mechanisms, adversaries may be able to gain complete control over the SCADA systems. Because of the user interface of the SCADA system, the adversary may have access to documentation and procedures for emergencies and change control, and this information may be enough for the adversary to control the system without any previous knowledge of how the underlying SCADA protocols work. Insiders with access to the control center may be able to compromise servers in the SCADA system and change their data. Adversaries without physical access to the systems may be able to exploit vulnerabilities in the network communication that is used between RTUs and the control center to get similar system access as an insider [17].

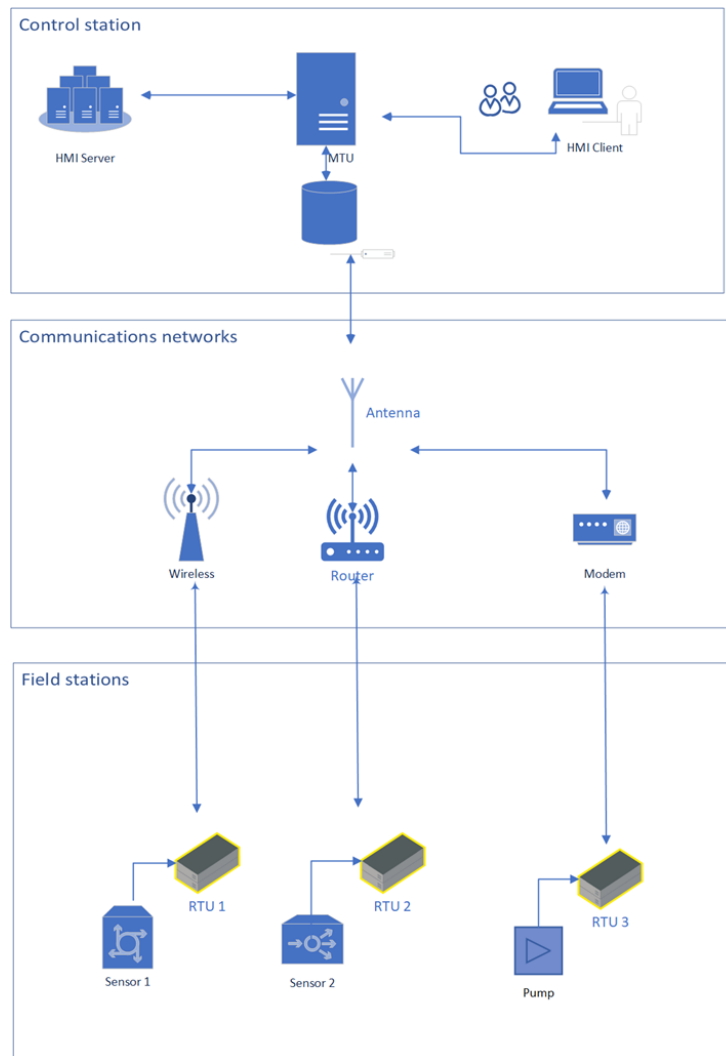


Figure 5: Typical infrastructure of a SCADA system

Both the Access Control Systems using electronic locks and the Human Presence Detection (HPD) using WiFi signals are tools that can be found in the STOP-IT toolbox, and can be used to protect systems in restricted areas. The Fine-grained Cyber Access Control Tool (FCAC), which is another tool from the STOP-IT toolbox, can be used to perform access control for both cyber and physical entities. For instance, if a user tries to access a database, the FCAC will check if the user has the right to access it and grant permission to access the database if the user has the right permission.

A Local Area Network is used for the communication between local processors and sensors and actuators. Analogue or digital signals are transmitted along relatively short cables or wireless connections [18]. Due to the geographical area of the water systems, a Wide Area Network (WAN) is used for the communication between local processors and the control center. To learn more about these technologies, have a look at the course module on communication systems.

The main priority of SCADA is to ensure safe, real-time operation, and to do so, the system must be constantly available. However, in the IT domain, confidentiality and integrity of the data has the highest priority. SCADA systems were initially created to work in closed environments, with no

connection to outside networks. The connection between SCADA and IT exposes SCADA components with barely any security mechanisms to the Internet, making it possible for adversaries to exploit vulnerabilities in IT to get access to the SCADA systems. Luijff [18] provides eight good practices for the technical management of SCADA systems. One of them is defence in depth. Defence in depth is provided by implementing security mechanisms between all layers of the SCADA system and the public and office network. If an adversary breaks one security measure it does not allow for uncontrolled access to the SCADA systems and network. The course module on communication technology explains defence in depth by giving a description of the Purdue model. In addition to defence in depth, operators should also consider horizontal segmentation of the system to ensure that systems that do not need to communicate cannot reach each other directly [3].

Cyber-attacks that target SCADA systems can be performed on hardware, software or on the communication links that connect the SCADA systems. Description of cyber-attacks targeting the communication link that connects the SCADA systems can be found in the course module on communication technologies. Software packages run on the hardware components and are used to collect data or send control commands to a process. A PLC is an example of a hardware component with software that enables it to automatically execute control activities [11].

The SCADA system for water systems is distributed and covers a wide geographical area. Sensors are dispersed throughout the water system, and if an adversary manages to get direct physical access to a sensor, such as a water level sensor, he or she may damage, manipulate, or even replace the sensor. The PLC may thus make decisions based on wrong input values, which may lead to deception or Denial of Service (DoS) [12]. A DoS attack is an attack where the wireless access is blocked because of adversaries sending large amounts of messages to reduce the network capacity or to shut down the network [3].

If an adversary manages to attack the communication links between the components in the SCADA system, he or she may manage to get direct control of a PLC in the network. Depending on the level of control gained, the adversary may perform a DoS attack against the PLC, manipulate the control logic of the PLC, or even provide the SCADA with incorrect data [12].

Cyber-attacks that target SCADA systems can be classified as either deception attacks or denial-of-service (DoS) attacks [13]. When performing a deception attack, the adversary sends false information from sensors or controllers. The false information may be incorrect measurements, incorrect timestamp of the measurement, or incorrect sender identification. Such attacks may be launched by obtaining secret keys used by the devices, or by compromising some of the sensors or controllers [14].

Evaluation questions

1.) What is the main priority of the SCADA system?

<input type="checkbox"/>	A: To ensure a secure, real-time operation
<input checked="" type="checkbox"/>	B: To ensure a safe, real-time operation
<input type="checkbox"/>	C: To ensure a safe operation, system delay is accepted
<input type="checkbox"/>	D: To ensure a secure operation, system delay is accepted

Note: B is correct

2.) How can the answer to question 1.) be fulfilled?

<input type="checkbox"/>	A: The system must be available during the day
<input type="checkbox"/>	B: The system must be available most of the time
<input checked="" type="checkbox"/>	C: The system must be constantly available
<input type="checkbox"/>	D: The system must be constantly available, but system delay is accepted

Note: C is correct

3.) How can defense in depth be provided?

.....

Note: SCADA system layers

4.) How does defense in depth work?

.....

Note: How does it protect the system?

5.) What may be the consequences of an adversary managing to attack the communication links between the components in the SCADA system?

.....

A.4. Sample material for chapter: Communication technologies for water

Topic

The topic covered in this course module is on **communication technologies** in (smart) water systems, as well as their main risks.

Note: This module is based on material from STOP-IT [5].

Goals

- Explain, in very simple terms, key communication technologies associated with (smart) water systems (wired and wireless), e.g., ethernet, WAN-based networks, servers etc.
- Explain the **risks** associated with these technologies. Emphasize the links with STOP-IT tools (e.g., mention jamming attacks).

Instructor pre-requisites

This topic should be taught by an instructor with competence in computer network engineering.

Text

In this module we will describe some of the communication technologies commonly used in (smart) water systems, the major problems associated with them, and give a link to some of the tools available in the STOP-IT toolbox.

Supervisory Control and Data Acquisition (SCADA) systems are Operational Technology (OT) used to monitor and control processes in the water systems. More information on the SCADA-systems used in water distribution, can be found in the course module on information systems (see A.3).

The water system covers a wide geographical area, and SCADA systems, therefore, use a communication network to allow system components to communicate. A Local Area Network is used for the communication between local processors and sensors and actuators and is typically the network organizations use to connect computers within a building. Analogue or digital signals are transmitted along relatively short cables or wireless connections [18]. Due to the geographical area of the water systems, a Wide Area Network (WAN) is used for the communication between local processors and the control center. A WAN is simply a network that covers a wide geographical area, and one example is the Internet. A WAN may also be the connection between several Local Area Networks. The STOP-IT toolbox provides a tool called Real-Time Sensor Data Protection, which applies blockchain schemes to protect the integrity of all the data generated during a critical infrastructure operation against both intentional attacks and malfunctions. This tool requires sensor data and the identification of the device that generated the data to be stored in the Cloud or in an alternative storage system.

Communication protocols, such as the Internet Protocol (IP), are used for communication between components in the SCADA system, for instance between the SCADA control center, which is the core of the SCADA system, and the local processors such as Remote Terminal Units (RTU) and Programmable Logic Controllers (PLC)¹¹. Figure 6 illustrates the conceptional layers of the TCP/IP model. According to Schneider-Electric¹² the Transmission Control Protocol (TCP), which runs on top of IP, is the most commonly used protocol for SCADA systems. The protocol offers error correction and uses a method called flow control, which can determine when a packet needs to be re-sent. This method ensures that packets are delivered as the RTU requests the packet from the host until the whole packet is complete and is identical to its original. The user Datagram Protocol (UDP) is also used, however, this protocol creates the packets and sends them to the host without establishing a connection first, meaning that it does not provide any form for authentication [19].

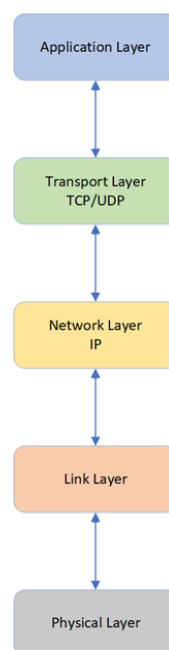


Figure 6: TCP/IP protocol stack

SCADA protocols, particularly those running on the top of transport protocols, such as TCP, have vulnerabilities that could be exploited by an attacker. The attacker may use methodologies such as injecting malformed packets to cause the receiving device to respond or communicate in inappropriate ways and result in the operator losing complete view or control of the control device [20]. According to Luijff [18], simple tests have shown that SCADA systems become stressed or even break down as soon as an unknown package is sent to them via the network.

¹¹ [Source: https://download.schneider-electric.com/files?p_enDocType=White+Paper&p_File_Name=998-2095-04-12-12AR0_EN.PDF&p_Doc_Ref=998-2095-04-12-12AR0_EN].

¹²[Source: https://download.schneider-electric.com/files?p_enDocType=White+Paper&p_File_Name=998-2095-04-12-12AR0_EN.PDF&p_Doc_Ref=998-2095-04-12-12AR0_EN]

Currently, SCADA runs applications on Windows or Linux and uses Internet Protocols (TCP/IP) for the transfer of data. The vulnerabilities of those systems and protocols are known by hackers all over the world and can be exploited using various toolkits [18]. There are several attacks that specifically exploits the implementation of TCP/IP protocols in Windows. However, as the systems operate in real time, it may be difficult for the operator to patch the systems right away [20].

The Modicon Communication Bus (Modbus) protocol is an application layer messaging protocol that operates in a master-slave mode. The master, for instance the HMI, initiates communication by sending query messages to the slaves, such as PLCs. The PLCs only respond to the queries, and do not initiate any communication with the master. The query may either be read or write [21]. There are two versions of Modbus, Modbus RTU and Modbus TCP. Modbus TCP is simply Modbus RTU implemented using TCP/IP over Ethernet¹³. Negi et al. [22] have performed penetration tests in order to reveal vulnerabilities in Modbus-TCP that could be used to attack PLCs. During the tests, they discovered several vulnerabilities. For instance, it is possible for an attacker to flood the server with specially crafted Modbus-TCP packets that leads to a Denial of Service (DoS) attack, preventing legitimate users from accessing the device – for instance preventing the HMI from requesting values from a PLC. Another thing that was discovered during the tests was that the Modbus-TCP plaintext communication reveals all the information about the addresses of the input/output for the PLC, which makes it easier for a malware to target specific addresses and attack the system. Due to the plaintext communication and no authentication mechanism built into the protocol, the PLC can be compromised by a Man-in-the-Middle attack. During the tests, they were also able to run a malicious script to execute commands. The Modbus-TCP server running in the PLC accepts connection without authorization and serves the client by executing commands.

The Distributed Network Protocol 3 (DNP3) is a protocol that enables the communication between the SCADA control center and the RTUs and PLCs, which is referred to as outstations or local processors [20]. DNP3 was designed to address the security gap of Modbus [23]. One benefit with this protocol is that it allows for communication between equipment from different vendors. The DNP3 packets are encapsulated by TCP or UDP packets and uses several standardized data formats and support time-stamped data, making the data transmission reliable and efficient¹⁴. Many manufactures choose to use this protocol due to its small memory consumption. When using DNP3 it is possible to reset and reconfigure the local processors by forcing them to turn off, wait for a while, and then turn it back on again. An attacker may take advantage of this property to cause delay in outstations before they accept requests again [20].

DNP3 unsolicited messages are messages that allow the RTU to send specific messages to the master station without getting a request first. These messages may be warnings or error messages that need immediate actions. When sending these messages, the RTU expects to receive a control message from the master station. An adversary may perform an attack by stopping the RTU from sending these messages. This type of attack may cause great damage to the system [24].

Another type of attack is the DNP3 data set manipulation attack. This attack exploits DNP3 messages that are sent over TCP/IP without using any form of transport layer encryption. An attacker may for

¹³ [Source: https://www.modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf]

¹⁴ [Source: https://www.westermo.com/-/media/Files/White-Papers/westermo_wp_industrial_protocols_and_weos_compatibility.pdf]

instance modify the content of the TCP message or replace the entire message with a new one. By modifying the TCP/IP header and DNP3 messages, the attacker can manipulate, control and redirect the DNP3 traffic and even modify the exchanged messages between the master and the RTU [24].

Jamming attacks are DoS attacks where an adversary interferes with the radio signal between two devices [25]. The goal of a jamming attack may be to overwhelm the system by sending transmissions faster than they could be handled, or by sending packets created to cause software errors on the targeted device. A jamming attack can be performed against SCADA communication technologies by jamming the traffic between a sensor and a PLC. The STOP-IT toolbox provides a jammer detector that secures the communication on the Wireless Sensor Network (WSN).

The Purdue model gives an overview of how the SCADA network can be organized and separated from the office network, using firewalls and a Demilitarized Zone (DMZ). An example of the model is illustrated in Figure 7. SCADA systems can be divided into three zones, the Operational Technology (OT) zone, the DMZ, and the Information Technology (IT) zone.

The OT zone is divided into four levels. The lowest level consists of the physical devices, such as sensors and actuators, that are to be controlled and monitored. The PLCs or RTUs control the processes of the physical devices and belong to Layer 1 of the model. Safety Instrumented Systems (SIS) are also implemented in this level. In Level 2, the SCADA system gathers and stores all the data that is shared by the lower levels. An HMI is used by operators to monitor and control the processes of the lower layers. System alarms are also provided to the operators through the HMI. The third level, the operation level, consists of applications, such as expert systems, maintenance systems and historian [26].

The DMZ zone controls and handles the data traffic between level three and four. Here, firewalls and other necessary equipment is used to ensure that level four does not have to directly access equipment in the lower layers. Equipment that is dedicated to ensuring ICT-security, such as malware monitoring applications and servers, equipment certificates that sends data to the lower levels are placed in this zone. The aim of this zone is to protect the entire OT network in such a way that all traffic to and from the OT network must go through nodes in the DMZ. The DMZ is isolated from both the OT system and external networks using firewalls [26].

The IT zone, or the Enterprise zone, is the office network. The applications in this zone may be reached from outside the zone through a firewall. Applications placed there, may for instance be applications for remote access to the systems in the OT zone [26].

Security measures are used to prevent adversaries from gaining access to the SCADA systems and networks through public or company networks. In the STOP-IT toolbox, there are several tools that can be used to protect the system. The Network Traffic Sensors and Analysers (NTSA) [60] is a STOP-IT tool that analyses the Netflow traffic data generated by routing and switching devices to detect anomalous behaviour in the traffic. All traffic considered abnormal will generate an alert to the STOP-IT XL-SIEM [61], which in turn will notify the STOP-IT RTAD [61].

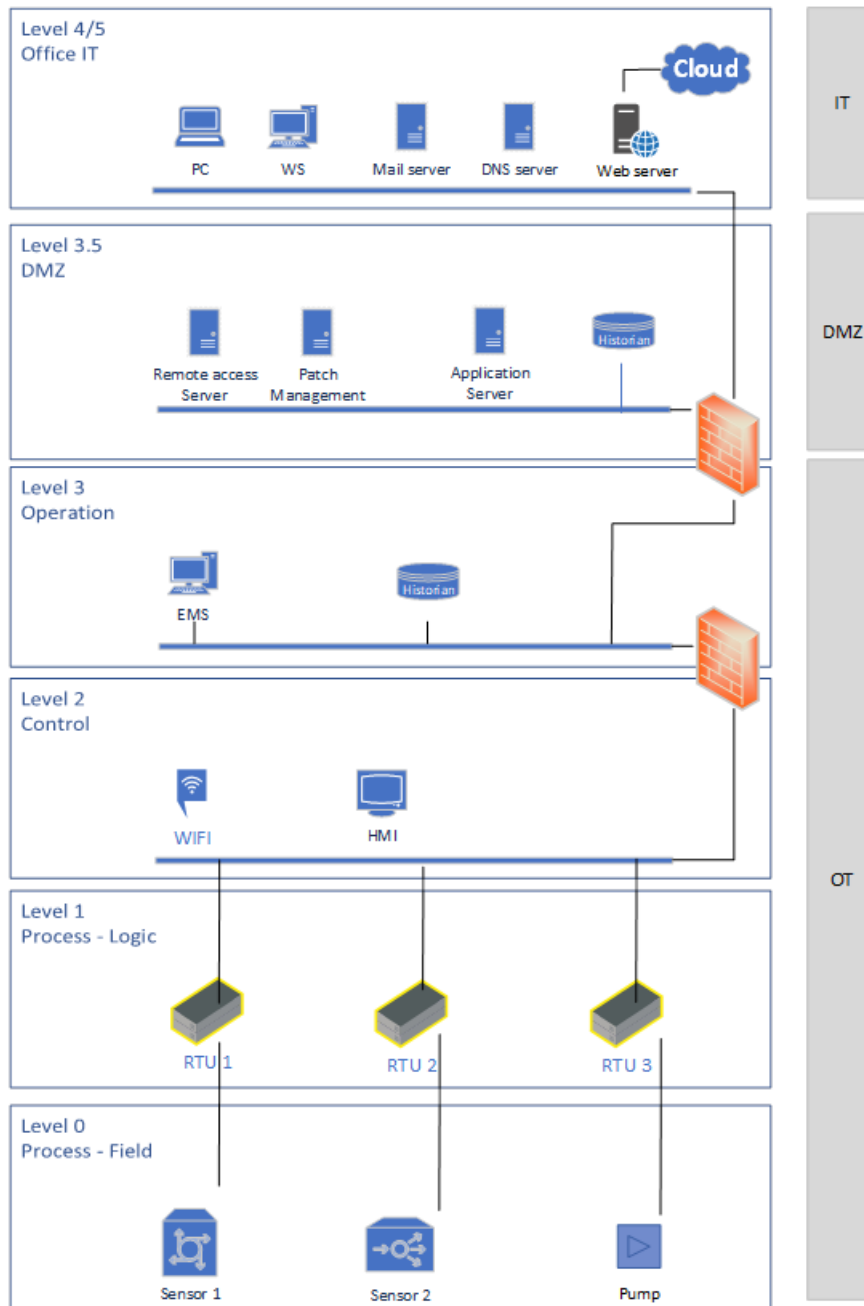


Figure 7: Purdue model

Evaluation questions

1.) What is a DoS attack?

<input checked="" type="checkbox"/>	A: An attack that blocks the wireless access
<input type="checkbox"/>	B: An attack that interferes with the frequency band by creating noise
<input type="checkbox"/>	C: It stands for distribution of service
<input type="checkbox"/>	D: It is an attack that tampers with the messages between two entities

Note: Answer A is correct

2.) Why is it difficult to patch Windows systems when new security updates are available?

<input type="checkbox"/>	A: Because employees think it takes too long, and don't want to wait
<input type="checkbox"/>	B: Because SCADA run applications on Windows, and to ensure a secure operation they need to be constantly available
<input checked="" type="checkbox"/>	C: Because SCADA run applications on Windows, and to ensure a safe operation they need to be constantly available
<input type="checkbox"/>	D: The patches are often expensive

Note: Answer C is correct

3.) What is a DMZ?

<input type="checkbox"/>	A: It is the office network zone
<input checked="" type="checkbox"/>	B: Demilitarized zone
<input checked="" type="checkbox"/>	C: Ensures that the levels above and underneath does not talk directly to each other
<input type="checkbox"/>	D: A zone that allows direct connection between SCADA components and the Internet

Note: Answers B and C are correct

4.) Mention some of the vulnerabilities that have been revealed in Modbus/TCP

.....

5.) What is a jamming attack?

.....

6.) Give a brief description of the Purdue model

.....

.....

.....

Note: DMZ, firewall, zones, levels

A.5. Sample material for: Risk management process for cyber-physical security

Topic

The topic treated in this course module is on the customization of the **risk management process**, described in general terms by the ISO Standard "Risk Management – Guidelines" (ISO 31000:2018) [59], to the objective of managing cyber-physical risk events in the water sector.

Goals

- To explain, in very simple terms, the risk management process described by the ISO 31000:2018.
- To present cyber-physical risk assessment tools for the water sector (also complementing the information provided in Module 3.2)
- To explain and support with an example, how the entire risk management approach can be applied having minimization of cyber-physical risk in the water sector as risk management objective.

Instructor pre-requisites

This topic should be taught by an instructor with competence in risk management and water network engineering.

Text

The risk management process described by the ISO 31000:2018 standard¹⁵

Risk Management is a Process of: Identifying Risk; Assessing and Prioritising Risk; Planning to Minimise Risk.

A simple risk management can boil down to only a few questions:

- What can go wrong/cause harm? Answering this question means performing risk identification.
- What is the chance of it happening? How bad? Can we accept it? Answering this question means performing risk analysis and evaluation.
- How can we minimise the chance of it happening? Or mitigate the consequences? Answering this question means performing risk treatment.

¹⁵ based on lecture at NTNU by Rita Ugarelli

ISO 31000 is a standard for risk management. First published in 2009, with the most current version (at the time of writing) being 2018, it describes a set of guidelines intended to streamline risk management for organizations. In ISO 31000, the focus is on best practice principles for implementing, maintaining, and improving a framework for risk management.

The ISO 31000 standard proposes an iterative framework to manage risk including the steps depicted in Figure 8. A short description of the scope of each step is described in the following sections.

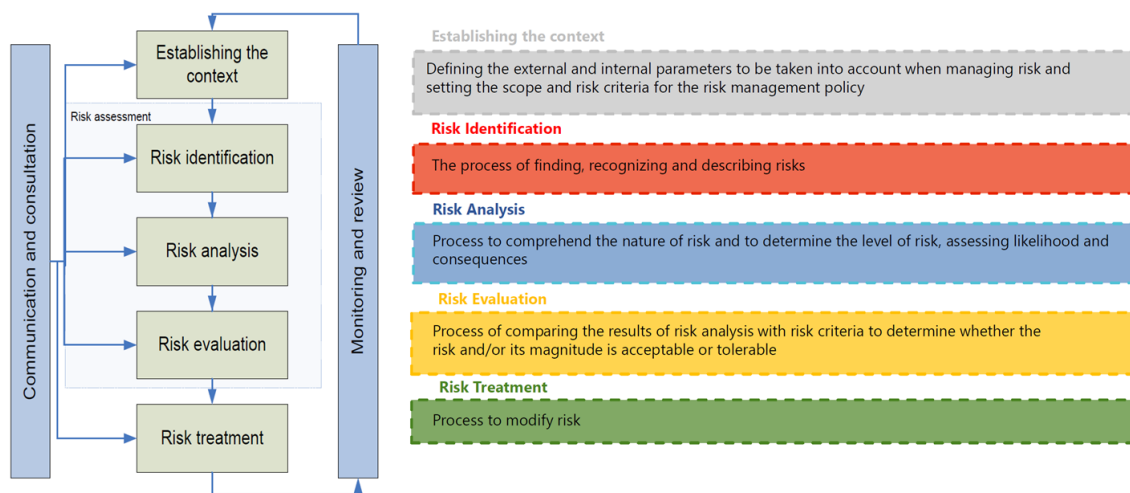


Figure 8: The risk management process described by the ISO 31000

Establishing the context for risk assessment

In each specific project, an overall problem definition is carried out prior to risk assessment. The context for the risk assessment has to be established before commencing the exercise of risk identification. Establishing the context defines the scope for the risk management process, and the primary objectives of the utility (of the service provided) and sets the criteria against which the risks will be assessed. The following sub-steps have to be performed:

Describe the system and its subsystems

A full, detailed and updated description of the system and, if existing, subsystems should be created (e.g. with flowcharts).

Define the scope for risk management

The scope of the risk management process should include the specification of:

- the objectives of the organization (e.g., for water utilities the objectives could be protection of public health, public safety and environment)
- parts of the organization (activities, processes, functions, projects, products, services or assets) where the risk management process will be applied;
- risk assessment methods to be applied.

Set risk acceptability criteria

The criteria to be used in the evaluation of the significance of risk must be defined in the light of the utility's objectives. Legal, regulatory or other type of formal requirements can impose some of the criteria. The following aspects should also be considered in criteria setting:

- nature and types of causes and consequences that can occur and how they will be measured;
- how likelihood is defined;
- timeframe of the likelihood and/or consequence(s);
- how the level of risk is defined;
- level at which risk becomes acceptable or tolerable;
- whether combinations of multiple risks should be considered; if so, how and which combinations should be considered.

Often due to data limitations a qualitative approach is used to define the levels of risk:

- In this case it is necessary to define a likelihood qualitative scale, the consequence dimensions qualitative scales and a risk matrix, as well as the criteria for risk evaluation for each risk level.
- These scales and criteria should be defined in close collaboration with the decision makers to ensure consistency with their values and strategies. The scales should be made before starting a risk identification phase, but this is not strict at all.

What is a risk matrix?

A risk matrix is a method that provides an approximation to a quantitative relation between Consequence (C) and Probability (P)

- Estimate a risk level of identified risk events
- Sets the risk criteria: the levels of acceptable risk
- Discriminate between three levels of risk associated to acceptance criteria
 - Low (acceptable)
 - Medium (Tolerable)
 - High (not acceptable)

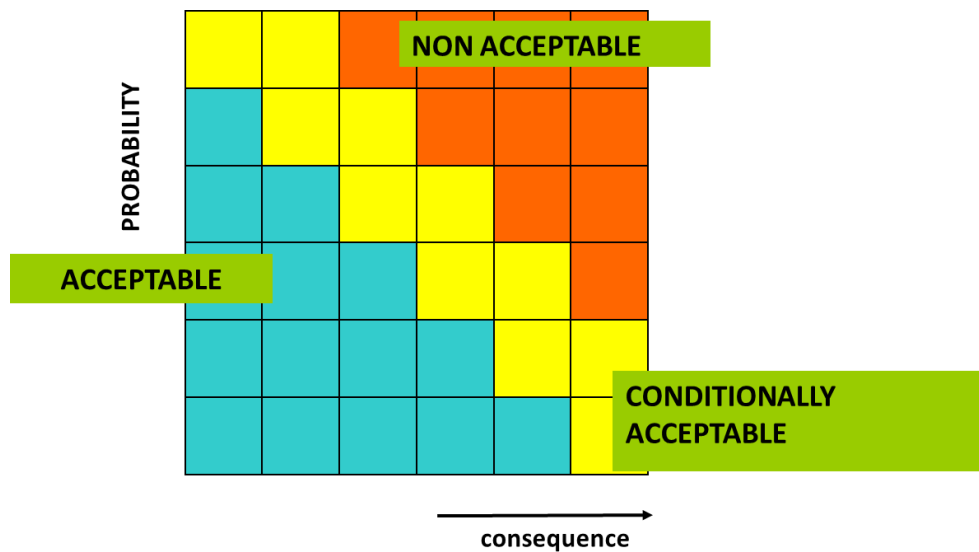


Figure 9: Example of risk matrix

Risk identification (RI)

This is the step resulting in a list of risk events. At this stage we only think of all what can go wrong, without assessing it.

There is no specific formula to describe a risk event; however, there is guidance provided in the ISO 31000:2009/2018 Risk management—Principles and guidelines that can help to better articulate risk events as a structured and concise explanation of what occurs in the event, usually including the pathway of the event.

An example of risk statement should reflect:

[Event that has an effect on objectives] **caused by** [cause/s] **resulting in** [consequence/s].

An alternative statement version is:

[Event that has an effect on objectives] **caused by** [cause/s]. **This may result in** [consequence/s].

Risk analysis (RA)

For each plausible event, the likelihood and the consequences should be estimated using the method selected at the step "establish the context"

Likelihood: chance of something happening, whether defined, measured or determined objectively or subjectively, qualitatively or quantitatively, and described using general terms or mathematically such as a **probability** or a **frequency** over a given time period. Probability is the measure of the chance of occurrence expressed as a number between 0 and 1, where 0 is impossibility and 1 is absolute certainty. In some languages **probability** is used with the same broad meaning.

Consequence: outcome of an **event** affecting objectives. An undesired event can lead to a range of consequences.

Analytically each event (E_i) is associated with its likelihood (l_i), but can be associated with more than one type of consequences (c_{ij}).

$$r_{ij} = f(l_i, c_{ij})$$

The resulting risk level is defined as r_{ij} .

Thus, each risk event can lead to more than one result in terms of risk, each for a different consequence dimension.

Risk evaluation (RE)

Risk evaluation involves comparing the levels of risk estimated during the risk analysis with the risk criteria established (possibly during the "establish the context phase").

Risk events can then be ranked in terms of severity.

The results are used to make decisions about future actions on: risks that need treatment; priorities for treatment actions.

Following frequency and consequences assessed in the risk analysis step, events are for instance put to the risk matrix and then compared with the risk evaluation criteria (Figure 10)

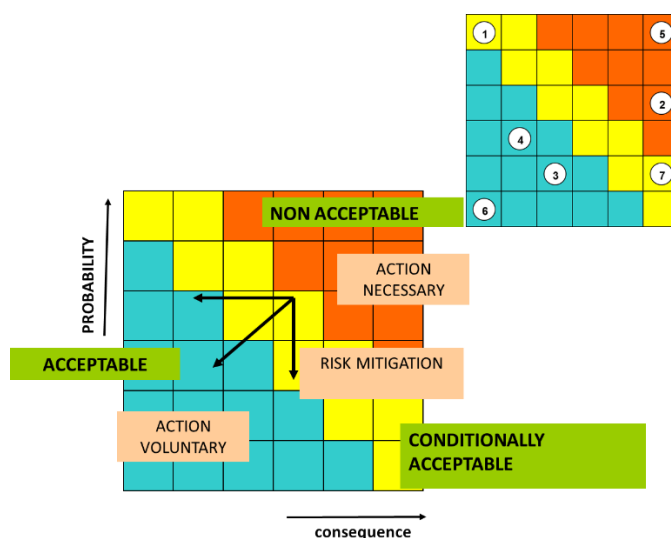


Figure 10: Use of risk matrix in the risk evaluation step: numbers 1-7 refer to analyzed risk events

Risk Treatment

The purpose of risk treatment is to select and implement options for addressing risk.

Risk treatment involves an iterative process of:

- formulating and selecting risk treatment options, often referred as Risk Reduction Measures (RRM). A list of possible RRM should be obtained for each event with a non-acceptable level of risk;
- planning and implementing risk treatment;
- assessing the effectiveness of that treatment. The task is to analyse the risk reduction "ability" of a single RRM in order to decide whether the measure is worthwhile or not;
- deciding whether the remaining risk is acceptable;
- if not acceptable, taking further treatment. Often, multiple measures might need to be considered and, given the usual limitations on resources to implement measures or simply because they might be redundant, it is necessary to compare and rank them by its effectiveness.

To rank the RRMs and select one for an event it is necessary to assess the effect in reducing risk for each RRM k , for a given event (E_i) in a specific consequence dimension (j). Basically, the effect of risk

reduction can be expressed as $\Delta r_{ijk} = r_{ij} - \overline{r_{ijk}}$

where the effect of the risk reduction (Δr_{ijk}) of the measure k is defined as the difference between the level of risk before treatment (r_{ij}) and the level of risk after treatment with the measure ($\overline{r_{ijk}}$)

So, **the main objective of risk treatment is to reduce risk to an acceptable level**, as visualized in Figure 11.

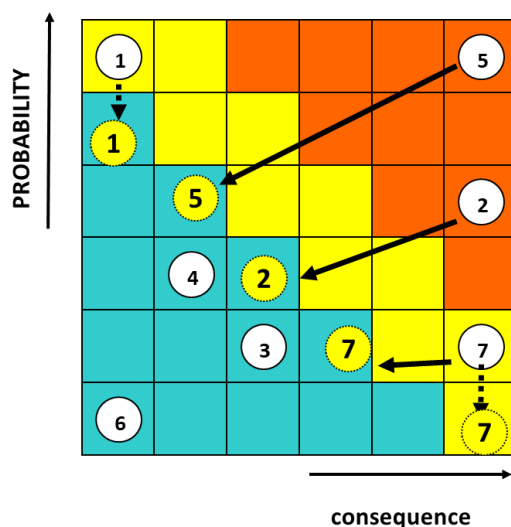


Figure 11: Lowering risk levels through adoption of RRMs

Monitoring and Review

The purpose of monitoring and review is to assure and improve the quality and effectiveness of process design, implementation and outcomes. Ongoing monitoring and periodic review of the risk management process and its outcomes should be a planned part of the risk management process, with responsibilities clearly defined.

Monitoring and review should take place in all stages of the process. Monitoring and review includes planning, gathering and analysing information, recording results and providing feedback.

Communication and reporting

The risk management process and its outcomes should be documented and reported through appropriate mechanisms. This step aims to:

- communicate risk management activities and outcomes across the organization;
- provide information for decision-making;
- improve risk management activities;
- assist interaction with stakeholders, including those with responsibility and accountability for risk management activities.

Cyber-physical risk assessment tools for the water sector

The tools and methods hereby suggested for the cyber-physical risk assessment for the water sector are based on the Risk Management Guide reported in the DWC deliverable D4.3 [27].

Risk Identification

A proper risk description should comprise four elements, namely sources, type of event, causes, and consequences. Information useful to support risk identification include expert knowledge and judgement, personal and organizational experiences, checklists, historical records, incident databases, previous risk registers, and reports from previous risk assessments. Identifying the events and their possible paths is an important and not straightforward step in the RMP. Each risk event has its causes and understanding them can significantly help in estimating the level of risk. The risk causes, type of threat and consequences are parts of each event path.

The Risk identification Database (RIDB), available at <https://risk-explorer.digital-water.city/event> identifies the type of threats, the sources of risk, the description of the events and the type of consequences produced.

The RIDB of the DWC project builds on the approach provided by the STOP-IT project which focused on cyber-physical attacks in water supply systems. The events included in the RIDB should be considered as individual "building blocks" from which the complex risk scenarios can be derived by their combination. Therefore, the RIDB does not include events generated by the combination of multiple risks. The sentence's structure is the same for each record to ensure consistency, as shown in Figure 12.

A generates a **B** threat causing a **C** of the **D** of the **E** which affects **F** and might lead to a **G** issue

Type of source	Type of threat	Type of event	Supporting asset	Composite asset	Primary asset	Consequence
A	B	C	D	E	F	G

Figure 12: Records structure in the RIDB

Where:

- A: Type of risk source;
- B: Type of threat;

- C: Type of event;
- D: Supporting asset (involved specific element);
- E: Composite asset (involved solution);
- F: Primary asset (involved infrastructure);
- G: Type of consequence.

The user of the database can create a new event in the RIDB if the risk event of interest is not included, maintaining the same structure in each record.

Risk Analysis

For risk analysis, there are three types of methods used for determining the level of risk, namely qualitative, semi-quantitative, and quantitative methods.

Qualitative Methods. This type of approach is often adopted for decision-making based mainly on expert judgment, experience, and intuition. These methods can be used when the level of risk is low and does not warrant the time and resources necessary for making an extensive analysis. These methods are also used when the numerical data available are not adequate for a more computational and quantitative analysis, so it would serve as the basis for a subsequent and more detailed analysis. The qualitative methods include brainstorming, questionnaires and interviews, evaluation for multidisciplinary groups, judgment of specialists and experts, etc.

Semi-Quantitative Methods. In this type of approach, classifications and scores based on empirical formulas are usually adopted, with calculations targeted on retrieving the ranges of likelihood and consequence of a certain risk event. The classifications are shown in relation to an appropriate scale for calculating the level of risk. High attention should be given with respect to the adopted scale, in order to avoid misunderstandings or misinterpretations of the results of the calculation.

Quantitative Methods. This type of approach allows to assign non-discrete values of loss to the various risks identified, enabling the calculation of the level of risk for several scenarios of attack. The quantitative methods include analysis of likelihood and consequences usually computed by multiple simulations (e.g., Monte Carlo simulations). The assessment of the consequences could be expressed in terms of KPIs related to different dimensions (finance, health, reputation, environment, etc.), depending on the nature of the risk in which the organization is interested. Given a digital twin of the system, stress-testing can be adopted as a method to compute potential impact given the cyber-attack resulted successful. On the other hand, if historical data are available, the probability of a successful cyber-attack could be computed based on the recognized past malicious events. Combining the objectively estimated probabilities with the consequences computed through stress testing simulations, a quantitative risk analysis can be performed.

Consequence

In Risk Analysis, understanding how to model the risk event is the key, keeping in mind which data would be required in the analysis in relation to the risk criteria set a-priori, and which variables are the most relevant for the identified risk event (e.g., potential critical areas, number of affected individuals, etc.). During the RMP, the consequence assessed for the specific dimensions of impacts (e.g., economic, reduction of service, environmental, organizational resilience, etc.) could be quantified through Key Performance Indicators (KPI).

In the water infrastructure domain, a methodology which involves the stress-testing of drinking water supply systems has been developed in the STOP-IT project through the **RAET** (Risk Assessment and Evaluation Toolkit). The stress-testing platform (STP) integrated in RAET can simulate both physical and cyber sub-systems coupling the simulation environment for the physical layer to an emulation environment able to model the cyber layer of the water system control and communication infrastructure (e.g., from SCADA to PLCs to monitoring), where cyber protection solutions will be implemented, and cyber-attacks attempted.

RAET builds on the risk management process described by the ISO 31000:2009/2018 (defining the steps of risk identification, analysis, evaluation, and treatment) and adapts its steps and methodologies to serve the needs and security scopes of cyber-physical security. For the assessment of consequences of a cyber-physical attack, RAET covers the following tools:

- The Asset Vulnerability Assessment to risk events Tool (AVAT) to show the criticality of each element in the water network, using vulnerability metrics, such as the Link Criticality Index, defined as the number of disconnected nodes resulted from an element outage. The tool helps to handle the complexity of water distribution networks, where usually it is not trivial to gain knowledge on the location of the vulnerable components. The tool can score the system vulnerability for different configurations of the network, providing a ranking of pipes based on the potential impact on the system that each single pipe would have if its failure has occurred.
- A scenario planner designed to assist the user in creating the scenarios of attack to be examined; it is supported by the STOP-IT RIDB from which potential generic risk events to be analyzed can be selected and the designed STOP-IT Fault Trees (FTs) to navigate through multiple paths from threats to events; the built scenarios of attack can then be further examined and simulated within the Stress Testing Platform or any other user selected model.
- A Stress Testing Platform that can simulate both physical and cyber sub-systems coupling the simulation environment for the physical layer to an emulation environment able to model the cyber layer of the water system control and communication infrastructure (e.g., from SCADA to PLCs to monitoring), where cyber protection solutions will be implemented, and cyber-attacks attempted. The platform allows to analyze for example the effects of introducing malware to the supervisory system and trace these effects to Key Performance Indicators (KPIs).

Probability

The evaluation of the probabilities of successful cyber-attacks on CIs could be challenging when a new digital solution has been recently developed and/or no historical records about past events are available. A semi-quantitative approach for vulnerability and exposure assessment of systems to cyber-physical attacks, developed with **InfraRisk-CP** [63] in the STOP-IT project, can be adopted.

Probabilities of a successful attack might be evaluated based on a structured subjective assessment which should consider expert judgments about multiple aspects such as the attractiveness of the assets or the hacking capabilities of the attacker with respect to vulnerable parts of the existing IT systems. In InfraRisk-CP, the following approach is used to assess the frequency of a successful attack:

1. To find the frequency of an attack attempt (sometimes referred to as likelihood of threat happening) a set of questions is provided.
2. For each question there is a predefined list of answers, where each answer is associated with a score.
3. The scores are aggregated to give a total score for the frequency of an attempt.
4. To transform the score to a frequency number a low value f_L and a high value f_H are defined. f_L represents the frequency of an attack attempt if all scores for the attack attempt questions have the lowest possible values, and f_H represents the frequency of an attack attempt if all scores have the highest possible values.
5. To find the probability of the success of an attack attempt (sometimes referred to as likelihood of threat succeeding) another set of questions is provided.
6. For each of these questions there is also a predefined list of answers, where each answer is associated with a score.
7. To transform the score to a probability number a low value p_L and a high value p_H are defined. p_L represents the probability of a successful attack attempt if all scores have the lowest possible values, and p_H represents the probability if all scores have the highest possible values.
8. To find the frequency of a successful attack attempt, the frequency of an attack attempt is multiplied with the probability of success.

For assessing frequencies of cyber or physical attacks, a list of questions is provided, where scores are obtained for each sub question (S_1 , S_2 , etc.). Scores are aggregated according to the formulas described in the InfraRisk-CP manual [63]. If no information is available for a given question, a score of 3 is given. If a question is not considered relevant, the score is excluded from the aggregation.

Risk Treatment

The DWC project provides a database called RRMD (available at <https://risk-explorer.digital-water.city/measures>), where several risk reduction measures were gathered and associated to related risks events of the RIDB. Similarly to the RIDB, the RRMD was firstly developed under the project STOP-IT within the domain of water supply systems, then the database has been populated with generally described Risk Reduction Measures (RRM) to extend the applicability to other systems, also beyond the scope of DWC project. This ensures the implementation of the listed measures in a large variety of cases. A many-to-many relationships between risk events of RIDB and measures of RRMD can be realized, thus an event of the RIDB may be associated to several suitable measures of the RRMD. On the other hand, a measure from the RRMD may address several risks documented in the RIDB.

Since the RRMD is not and cannot be an exhaustive list of all possible RRM, it shall not supply a fully prepared and formulated plan for Risk Treatment, but rather show to the user options on how existing risks could be treated by choosing and implementing one or several measures. Thus, it is important to enable future users of the tool to populate the database with additional measures. Only by keeping the RRMD a “living register” its practical value can be ensured also in the future, also with respect to incoming cyber-physical threats of critical infrastructures, so the users may contribute by adding new relevant measures in the database

The risk management process applied to the water sector protection against C-P threats

Establishing the context for risk assessment

Describe the system and its subsystems

The WWTP in the city of Copenhagen, managed by BIOFOS, was selected as case study. One of the BIOFOS's main objective is the reduction of the pollution of the environment deriving from the treatment activities. Risk management is applied to deal with cyber-physical attacks or incidents which may damage the environment when the infrastructure of provided services are not sufficiently protected. The WWTP is characterized by four lines where all the biological treatment steps are run in parallel. The capacity of each line is equal to 2.500 m³/h, for a total treatment capacity of 10000 m³/h. To cover the high peaks of inflow due to the rain events, the WWTP is equipped with equalization tanks with a total volume of 44000 m³. In Table 2, the characteristics of the system are reported.

Table 2: Characterization of the considered water system

Characteristics of the System	Values	Units
Number of Treatment Lines of the WWTP	4	[-]
Capacity of each Treatment Line	2500	[m ³ /h]
Total Volume of the Equalization Tanks	44000	[m ³]
Time to restore the Line under Maintenance	24	[h]

Define the scope for risk management: Minimization of environmental risk while adopting a digital solution for the improvement at WWTP of maintenance schedules.

During and just after the rain events, given a certain hour of the day, when the actual inflow value is much higher than the expected wastewater flow, for BIOFOS the inflow can be considered enough diluted. Thus, a biological treatment could be by-passed directly to the receiving water body, without any significant environmental consequences. This concept has been derived from common design criteria of CSO (Combined Sewer Overflow) devices in Europe. Specifically, a dilution coefficient r - given in (1) with a value ranging from 3 to 6 is considered as the minimum critical value r_c under which all the inflows should be properly treated.

$$r = \frac{Q_{ww} + Q_r}{Q_{ww}} \quad (1)$$

where Q_r is the inflow generated by the rain and Q_{ww} is the contribution of wastewater flow. Obviously, in the stage of CSO devices design, the adoption of lower values of the critical dilution coefficient r_c leads to an increase of CSOs and a consequent decrease of biologically treated volumes. For the case study, the risk criteria have been defined considering the critical dilution coefficient r_c equal to 3. When the system has reduced capacity due to maintenance operations, the equalization tanks might not be enough to cover water overloads, since with reduced capacity the plant might not be able to treat all the wastewater characterized by dilution coefficients below the designed r_c . Knowing in advance the expected inflow of the next 48 hours, the web application developed within

the DWC project would help in planning the maintenance on each of the WWTP lines, without consequences in terms of undiluted and untreated overflows of wastewater, despite the reduced treatment capacity.

Usually, the maintenance is performed during dry weather because in that case there is not necessity of using all the four parallel treatment lines of the WWTP. The critical conditions should be identified during the phase of risk analysis through a stress-testing procedure, but it is already possible to hypothesize in advance that they are likely to occur if the attack is performed few hours before or during small-medium rain events. In this case, the operators might perform maintenance expecting dry weather conditions, but they will eventually experience higher inflow than expected because of the considered attack, thus this will potentially lead to not treat wastewater not sufficiently diluted, according to the designed r_c .

Set risk acceptability criteria

The water organisation aims at identifying the most hazardous scenarios of pollutant concentration, where eventually under cyber-attack, the polluted untreated overflows released in the environment might have a dilution coefficient r below 3. On top of the Risk Management steps, at least one risk criterion must be defined, according to the objectives of the involved organization, in relation to a selected KPI. For the case study, the yearly cubic meters related to biologically untreated and not diluted volumes (under the selected threshold of dilution coefficient) of wastewater were considered within a maximum duration of 24 hours per event. The constrain of 24 hours per event was considered because for the specific identified risk, the organization stated that an eventual emergency can be recovered within 24 hours.

Note that each considered gate valve of the WWTP is being maintained once every two years and given that there are four treatment lines, the maintenance on one of the gate valves is being executed twice per year in average, according to the WWTP manager.

The value of the selected KPI, expressed in m³, is the yearly maximum polluted cubic meters of wastewater, related to the worst event with a maximum duration of 24 hours. Based on the internal objectives of the organization, thresholds of levels of Risk have been defined, according to Table 3.

Table 3: Pre-defined level of Risk expressed in terms of the selected KPI

Low Risk	Medium Risk	High Risk
KPI ≤ 120	120 < KPI ≤ 1.200	KPI > 1.200

The threshold between medium and high risk has been set to 1.200 m³ of undiluted wastewater, corresponding to the estimated minimum wastewater inflow entering the WWTP for one hour during dry weather. The threshold between low and medium risk has been set to 120 m³ of undiluted wastewater, corresponding to the estimated minimum wastewater inflow entering the WWTP for ten minutes during dry weather.

Hence, risk criteria are defined through these thresholds which may trigger the Risk Treatment step, implying that the organization should consider risk mitigation measures if thresholds are overpassed with the already existing mitigation measures.

Risk identification

The considered digital solution allows a water utility to visualize the predictions of the flow entering in the WWTP with 48 hours in advance. Accurate rain forecasts are the input for a Machine Learning (ML) model which provides the timeseries of the mentioned flow, allowing optimized operations on the treatment process and improved schedules of maintenance. Specifically, the maintenance would be performed when dry time is expected. If a rain event is expected, the WWTP should work at full capacity because of the expected high loads to be treated. However, if the internal attacker modifies the visualized data to hide a rain event and the corresponding flow predictions, the water organization might start the programmed maintenance on one of the gate valves installed on the treatment lines. Therefore, if the internal attacker knows about the planned maintenance and manages to change the Web Interface visualization for the following 48 hours into a typical condition of dry weather, some issues may arise because of the unexpected inflow. According to BIOFOS, the recovery of full capacity would be restored in 24 hours.

Exploring the DWC RIDB, the water organization recognized a risk event generated by an internal attacker which could lead to quantity or quality issues on the effluent of the WWTP. Specifically, row 14 of RIDB is about the spoofing generated by an internal attacker of the web application for the visualization of WWTP inflow forecast. As reported in the database, considering the specific sentence structure related to each listed event, the risk is described in the following Figure 13, where the color code in the sentence depicts the corresponding elements of the adopted risk event’s syntax.

Internal attacker generates a **cyber** threat causing a **Spoofing** of the **Web application** of the **Web platform for integrated sewer and wastewater treatment plant control** which affects **Sewers or Wastewater treatment plant** and might lead to a **Quantity** issue

Type of source	Type of threat	Type of event	Supporting asset	Composite asset	Primary asset	Consequence
A	B	C	D	E	F	G

Figure 13: Selected event following the generic records structure of RIDB

Risk analysis: Estimate the consequences

The consequences are evaluated with a stress-testing approach for the case-study. A digital twin of the system has been adopted, considering primarily the capacities of treatment lines and equalization tanks. Stress-testing consisted in exploring the response of the digital twin of the system under the maintenance of one treatment line. All the available historical records of the inflow to the system have been considered as input and non-diluted and biologically untreated cubic meters of wastewater have been considered as output. In Figure 14, the available inflows of 2020 at the inlet to the WWTP are shown.

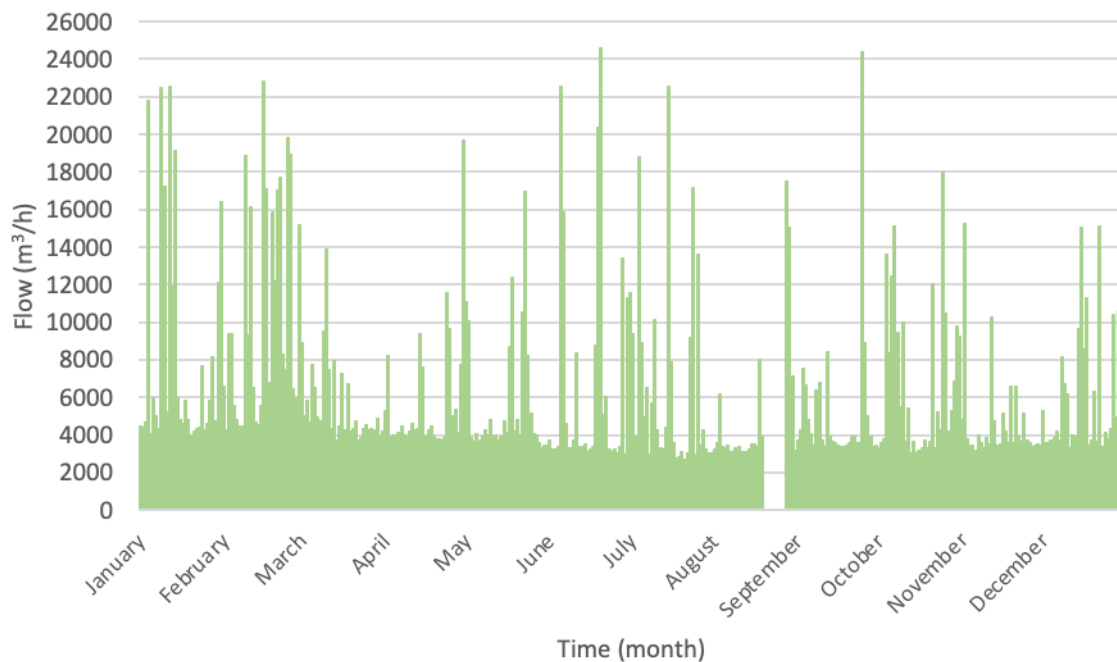


Figure 14: Measured flows at the entrance of the analysed Biofos WWTP in the year 2020

Specifically, the available data consisted in inflow values of the year 2020, with a resolution of one minute. The data shown in Figure 14 were firstly aggregated with hourly averages. Moreover, the missing data of part of the second half of August were excluded from the computations. According to the suggestions of the WWTP manager, one week of the first part of August (from 09.08.2020 to 16.08.2020) was considered as a reference for the hourly averages of wastewater flows Q_{ww} without any rain flows contribution. In general, during other periods, the inflow was higher because of frequent rain events and related high soil moisture levels which increased the run-off, arriving to the WWTP even with many hours of delay with respect to the time of actual rain precipitations. Having this selected week as a reference for the analyzed year, dilution coefficients were computed along the whole year. Concerning the critical dilution coefficient r_c , within the mentioned range of values between 3 and 6, as mentioned the lowest value equal to 3 was considered for the case study, however the level of accepted dilution can vary from case to case, according to the internal objectives of the organization. The equalization tanks are in general adopted to absorb additional loads which cannot be immediately treated by the operating treatment lines. For the simulation of the scenario of attack during the maintenance on one of the four treatment lines, the tanks were considered completely empty (thus with full capacity) on the 01.01.2020 and the water was considered as stored every time the flow at the entrance of the WWTP was greater than 7500 m³/h and released to the available treatment lines when the WWTP received less than 7500 m³/h. The available volumes in the process tanks of the WWTP were considered negligible with respect to the volume of the equalization tanks. The critical events along the year were identified, considering the organization could stop the consequences when issues last more than 24 hours. After running the simulations, the following critical inflow events are identified in the following Figure 15.

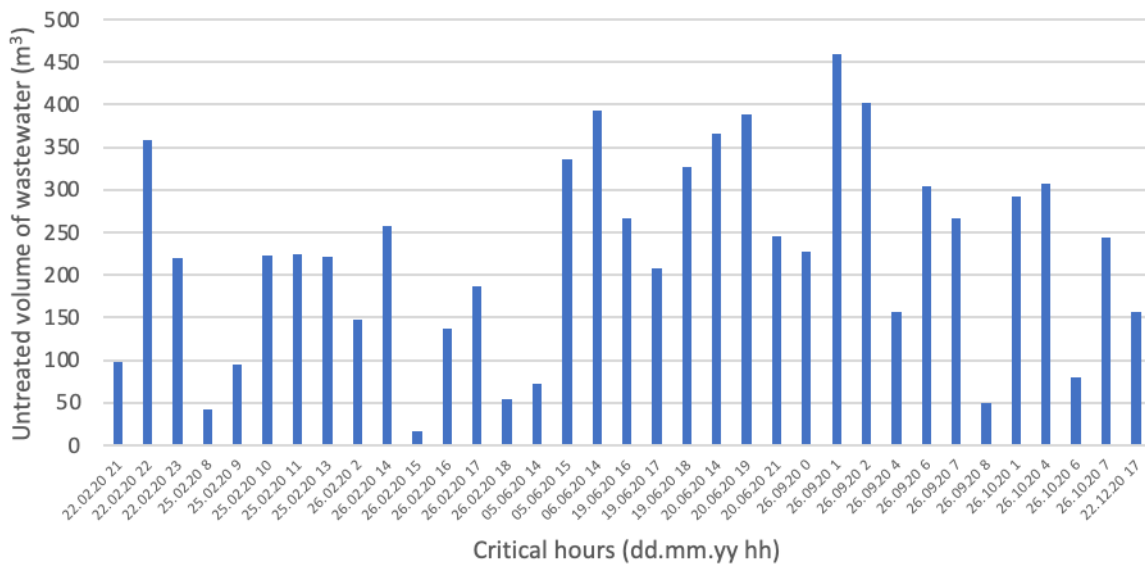


Figure 15: Critical events in 2020 identified with the stress-testing procedure.

Figure 15 reports the events of 22nd, 25th-26th of February (more than 24 hours), 5th-6th and 19th-20th (more than 24 hours) of June, 26th of September, 26th of October, and 22nd of December.

The maximum value of untreated volume of the year within a maximum duration of 24 hours provides the value of the selected KPI, in terms of consequences. Specifically, among the mentioned 7 events, the 26th of September 2020 is associated with the maximum untreated volume of wastewater within the same 24 hours, equal to 1.865 m³ and which represents the consequence KPI value for 2020.

Looking at the dates of the critical flow events, a retrospective assessment was performed with respect to the rain events of 2020 to better understand the involved risk factors. In Figure 16, the available rain events of 2020 for seven relevant stations (in the area of the studied catchment) of the Danish rain gauge network (SVK), expressed in µm/s and with a resolution of one minute, are reported.

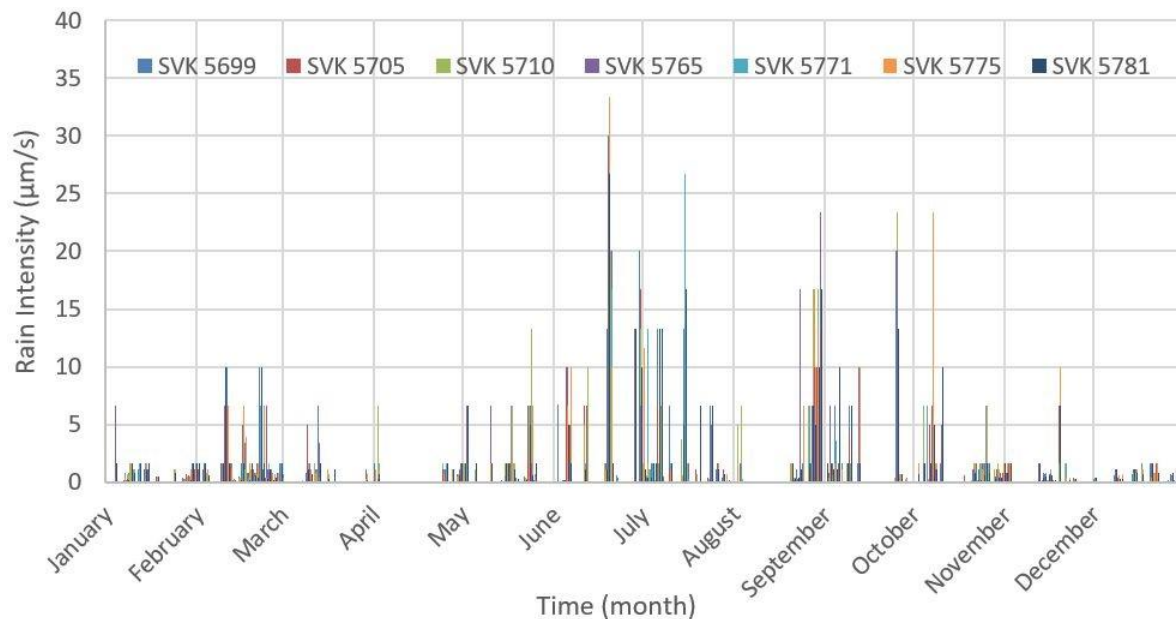


Figure 16: Rain data in 2020 of seven stations of the Danish rain gauge network

In terms of risk factors, it is important to highlight that the most dangerous conditions are not necessarily connected to the most extreme rain events (see, for instance, 22nd, 25th-26th of February, 5th-6th of June, 26th of October, and 22nd of December) mainly because the values of dilution coefficients would be higher than 3 for a large part of the event, so wastewater could be considered enough diluted, and the lack of biological treatment would not be a significant environmental issue; however, on the other hand, the equalization tanks might be more easily filled completely during extreme events. Moreover, extreme rains are more likely detected in advance through additional sources of information, thus for the attacker it would be more difficult to trick the WWTP operators and plan the attack during a scheduled maintenance just before a well-known expected extreme event.

Estimate the probabilities

A semi-quantitative approach for vulnerability and exposure assessment of systems to cyber-physical attacks, developed with InfraRisk-CP in the STOP-IT project, has been adopted for the case study. Probabilities of a successful attack might be evaluated based on a structured subjective assessment which should consider expert judgments about multiple aspects such as the attractiveness of the assets or the hacking capabilities of the attacker with respect to vulnerable parts of the existing IT systems. To evaluate probabilities or the frequency of a successful attack, the approach for cyber-attacks proposed in InfraRisk-CP from the STOP-IT project was considered.

It was assumed that the internal attacker is aware of the most favourable conditions for an attack, in terms of expected flows and planned maintenance in the system. Each considered gate valve of the WWTP is being maintained once every two years and given that there are four treatment lines, the maintenance on one of the gate valves is being executed twice per year on average, according to the WWTP manager.

Since the attack is internal, it is assumed that the attacker has the possibility to partially drive the schedules of the programmed maintenance at the same moment of the 7 events during the year which

produce consequences for the analyzed risk. The term "partially" is due to the extent of the attacker's will to drive the maintenance schedules and carry out the attack, which depends on the scores (S1-S15) of the InfraRisk-CP methodology. Thus, regardless of the attack attractiveness, assuming that over a period of 10 years the attacker would try at least one attempt, the frequency of attack spans from a minimum of once per ten year and a maximum of two times per years.

The frequency between these two extremes values derived by expert judgements for attempting the spoofing of the web application of the considered digital solution can indeed be estimated through InfraRisk-CP.

- f_L (Lowest attack frequency) = 0.1/year
- f_H (Highest attack frequency) = 2/year

The probability of success is mainly related to the capabilities of the attacker and to the security of the IT system. According to BIOFOS, the attacker is supposed to have good chances to penetrate the IT system since the considered attack is internal, but the actual value mainly depends on the capacities and the permissions already owned by the attacker. Specifically, if the attacker already has the full permission to the company's IT system, as an *administrative user* and the IT system has negligible protection in comparison with the attacker's capabilities, the probability is estimated to 100%. On the other hand, if the attacker's capabilities are negligible in comparison with the IT system and is only a *technical user* of the company, some effort on stealing the required credential of an *administrative user* would be needed, thus in this case the probability is estimated equal to 1%.

- p_L (Lowest attack success probability) = 1%
- p_H (Highest attack success probability) = 100%

In the following, the answers of BIOFOS to InfraRisk-CP questions are provided together with the calculations needed to estimate the frequency of successful attack.

1) How attractive it is to make an attempt to attack the water system, in terms of:

- Recognisability?

Answer: **S1 = 2 (low)**, due to the fact that there is no recognisability in affecting the wastewater treatment plants, power plants and distribution system are at a much higher risk.

- Symbolism?

Answer: **S2 = 1 (very low)**, this will likely not affect the citizens, but 'only' the environment, drinking water and distribution systems are at a much higher risk the wastewater treatment plants.

- Potential for economic profit (e.g., ransom)?

Answer: **S3 = 3 (medium)**, Organized crime does not specifically target wastewater treatment plants, but there is a medium risk.

- Potential for political profit?

Answer: **S4 = 1 (very low)**, Other utility sectors are at a much higher risk, electricity/power and drinking water utilities.

Note: *Recognisability* deals with attackers having a desired to be recognized within a community. Typically, this could be individual hackers. *Symbolism* could be relevant for terrorist groups which often have an objective to cause fear and uncertainty. Economic profit would relate to organized crime. Political issues could relate to foreign nations or political groups within one nation.

The scores S_1 , S_2 , S_3 , and S_4 could be seen as competing scores, and we let be a total attractiveness score $S_A = \max(S_1, S_2, S_3, S_4) + \Delta_A$. Here, $\Delta_A = 0.25 \ln n$, where n counts the number of scores equal the maximum score. Note that $\Delta_A = 0$ if the maximum score is 1 or 5 or the maximum score appears only once. In the analysed case **S_A is equal to 3.**

2) Level of Organizational issues, specifically regarding:

- Measures implemented towards insiders?

Answer: **$S_5 = 4$ (low)**, scarce employees' education regarding implemented IT security. User accounts for system access are in place, but no internal system to catch unsuccessful login/or hacking attempts.

- Quality of internal surveillance and intelligence systems?

Answer: **$S_6 = 4$ (low)**, no central system is implemented.

- Systematic preparedness exercises, investigation, and learning?

Answer: **$S_7 = 5$ (very low)**, never completed an exercise on the IT systems and infrastructure.

- Security focus in agreements with vendors and contractors?

Answer: **$S_8 = 4$ (low)**, vendors and contractors use to sign a confidentiality agreement regarding GDPR and information obtained during work/interaction with BIOFOS.

Note: For the organizational factors affecting the frequency of attack we calculate an average score: $S_o = (S_5 + S_6 + S_7 + S_8)/4$, so in the analysed case **S_o is equal to 4.25.**

3) Conditions affecting if an attacker will make an attack attempt for a specific component:

- How vulnerable the component seems from the attacker's point of view?

Answer: **$S_9 = 2$ (low)**, technical systems are behind the company firewall and a technical firewall that covers all the technical IT-systems. No administrative IT system user has direct access to the technical systems. A different technical username is required.

- Visible protective measures by the utility manager for the specific component.

Answer: **S10 = 2 (high)**, low, physical access to buildings and components is restricted. Alarm systems in buildings.

- How critical the component seems from the attacker's point of view?

Answer: **S11 = 2 (low)**, normal attackers does not have specific knowledge regarding the operations, equipment and control used at the wastewater treatment plant

- Accessibility of the particular component.

Answer: **S12 = 2 (low)**, all technical computer terminals are locked when not in use. Components (motors, gates) at the treatment plant cannot be operated locally when in automatic control mode.

- Attacker's capability vs required capability to make an attempt.

Answer: **S13 = 3 (medium)**, an attacker needs some skills to make an attempt, but it is possible.

- Attacker's available resources vs required resources.

Answer: **S14 = 3 (medium)**, an attacker needs good resources to make an attempt, but it is possible.

Note: For the conditions influencing willingness of an attacker to make an attempt an average score is also proposed: $S_w = (S_9 + S_{10} + S_{11} + S_{12} + S_{13} + S_{14})/6$, so in the analysed case S_w is equal to **2.33**.

4) Evidence with respect to possible attacks:

- How is the actual cyber security situation evaluated by the security authorities (police, intelligence, etc.)?

Answer: **S15 = 3 (medium)**, wastewater treatment plants are not the first in line for an attack, higher risk at power plants and power distribution and drinking water production and distribution.

- Evidence from internal surveillance for the specific attack (computerized monitoring tools).

This quantity is measured in terms of number of attack attempts per time unit, typically per year. Answer: **S16 not available**, main users cannot be currently detected, normal users would use workstation which are recognized; however, at the current time there are no evidence.

Note: To obtain a total normalized score for the likelihood of an attack, consider the average of S_A , S_o , S_w and S_{15} and standardize between 0 and 1:

$L = (S_A + S_O + S_W + S_{15} - 4) / (20-4)$, so in the analysed case **L is equal to 0.54**.

The frequency of an attack based on the influencing conditions is given by:

$$f = f_L \left(\frac{f_H}{f_L} \right)^L \quad (2)$$

The yearly frequency based on the assessment of conditions can be averaged with the observed frequency S16, if available. In the analysed case **f is equal to 0.5/year**.

For the probability assessment of a successful attack, another set of questions and related answers provided by BIOFOS are given in the following.

5) Likelihood of succeeding in an attempt:

- Attacker’s capability vs required capability to succeed in an attempt

Answer: **S17 = 4 (high)**, since the attacker is internal, but normally an attacker must overcome several firewalls and login to specific systems to succeed.

- Attacker’s available resources vs required resources to succeed in an attempt

Answer: **S18 = 4 (high)**, since the attacker is internal, but normally only highly trained attackers can access and penetrate the implemented security measures to gain access to technical systems.

- Explicit protective measures

Answer: **S19 = 2 (high)**, even if the attacker is internal because when using VPN access, encryption is used. Moreover, only VPN access from Danish IP addresses is allowed, a 2-step user verification for VPN access is adopted, and administrative IT user must login to VPN. To access the technical systems a technical user is allowed only via a VMware remote desktop, no direct server access. Finally, there are regular software updates of firewall, antivirus tools, clients, servers for both administrative and technical systems.

To obtain a probability measure for success of the attack, a standardised score is calculated in the interval from 0 to 1, with $Q = (S_{17} + S_{18} + S_{19} + S_6 + S_7 - 5) / 20$, so in the analysed case **Q is equal to 0.7**.

Note that in this score two of the organizational conditions are included. It could be argued that this gives “double counting”, but after the normalization, this is considered not to be an issue. The probability of a successful attack is given by:

$$p = p_L \left(\frac{p_H}{p_L} \right)^Q \quad (3)$$

In the analysed case **p is equal to 0.25**.

The frequency of a successful attack is given by:

$$f_A = f \times p \tag{4}$$

This frequency is expressed as a certain value per year, so in the analysed case f_A is equal to **0.125/year**, slightly more than once per ten years.

Risk Evaluation

With the stress-testing procedure applied to the system under a number of different configurations and based on the selected KPIs, it is possible to derive the conditions which may lead to the most serious consequences on the considered CI. After having combined the estimated consequences with their probability of occurrence, the risk levels are assessed by comparing the risk values with the risk criteria. The definition of the level of risk can lead to the optimal selection of the risk reduction measures to be implemented within the Risk Treatment phase.

The Risk Evaluation was derived from the previous phase of Risk Analysis. Specifically, the worst event of 2020 which may cause environmental issues was considered for the consequence assessment, in terms of the selected KPI, equal to 1.865 m³/year.

The probability of a successful attack per year is equal to 0.125, and it was computed through the approach which follows the InfraRisk-CP's methodology. Multiplying the consequences expressed in terms of the KPI with the probability, the risk is finally evaluated as **Medium Risk** as shown in Table 4, by comparing the results with level of risk defined in Table 3, since the risk is associated to the risk actual value (KPI) equal to 233 m³/year.

Table 4: Identification of the level of risk by comparing results with targets values

Low Risk	Medium Risk	High Risk
KPI ≤ 120 No Low Risk: 233 > 120	120 < KPI ≤ 1.200 Medium Risk: 120 < 233 ≤ 1.200	KPI > 1.200 No High Risk: 233 < 1.200

Based on this evaluation, different risk reduction measures could be adopted in the next phase of risk management, consistently to the actual level of risk.

Risk Treatment

The RRMD was adopted to explore different risk reduction measures alternatives for the case study. By exploring the RRMD the following categories of RRM have been selected by the user as potentially relevant to their case:

- Implementation of IT security systems
- Implementation of training procedure for the employees
- Increase of the volume of the equalization tank

In the first two cases, the risk might be prevented (probabilities of the occurrence of a successful attack are lowered), while in the third case the impact might be mitigated (the same probabilities hold but

the attack might have less impact on the system, in terms of the selected KPI). The suggested complete list in the RRMD connected to the selected risk event is shown in Figure 16.

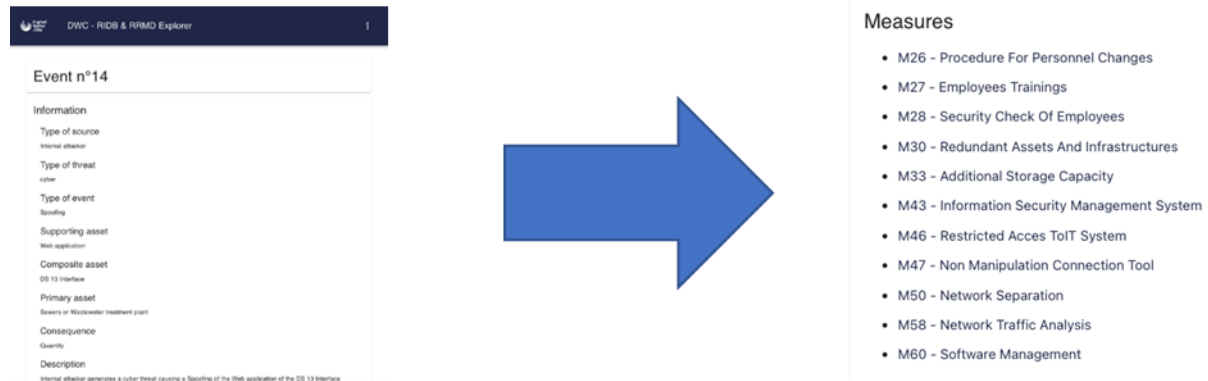


Figure 17: Risk reduction measures in the RRMD associated with the identified risk in the RIDB

The value of decreased probability of a successful attack when the first two types of measures are implemented are site-specific and depends on the current level of employees training and protection of the IT system.

In InfraRisk-CP, if all the quantities which depend on the level of protection of the organization (S5-S6-S7-S8-S17-S18-S19) are raised to their best score, the obtained estimation of the probability of a successful attack is equal to 0.017, leading to significant reduction of the risk, since it would be decreased almost to the 14% of the original value of risk, i.e., 33 m³, correspondent to **Low Risk**, according to the adopted risk criteria. On the other hand, if the organization implements actions to decrease the consequences (e.g., obtained with M33 of RRMD, thus for instance by doubling the volume of the equalization tanks), the associated KPI could be computed through the same procedure of stress testing, described in the risk analysis part. By adopting a doubled storage volume in the equalization tanks (88000 m³ of storage) and by following the same procedure described for the consequence assessment in the Risk Analysis paragraph, the hypothetical untreated polluted overflows of 2020 were computed and are reported in Figure 18.

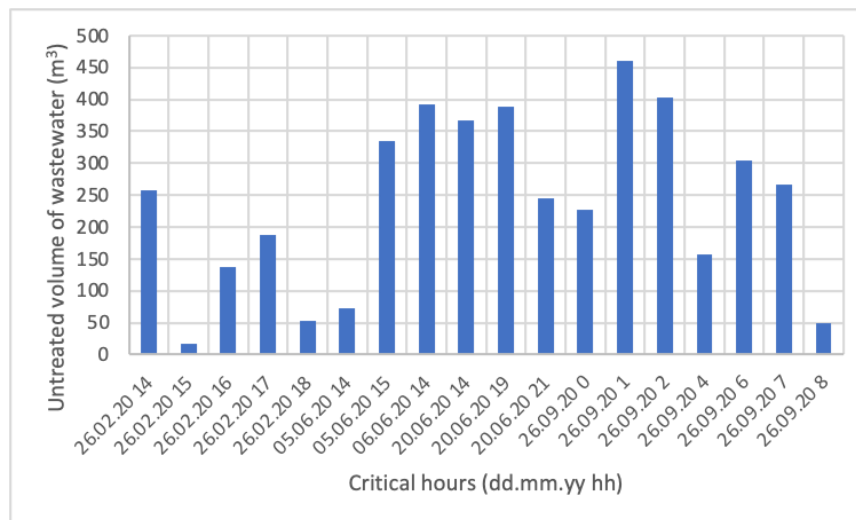


Figure 18: Critical hourly values in 2020 of wastewater without any rain contribution in the case of doubling the volume of the equalization tanks as risk reduction measure

Considering the definition of the selected KPI, its new value would result the same as before the implementation of the equalization tank, i.e., equal to 1.865 m³ in terms of consequences and 233 m³ in terms of risk, showing how important is to take decisions based on a structured RMP. In fact, even if the yearly untreated volume is globally reduced, the yearly maximum event, relevant to the performed risk assessment, does not report any significant improvement connected to this mitigation measure. On the contrary, if an additional treatment line with a capacity of 2.500 m³/h is installed (M30 in the RRMD), no negative consequences would be reported, since all the yearly inflow can be treated during the eventual considered attack. In this case, the water organization can rise its level of standards, for instance by increasing the minimum level of critical dilution coefficient. A value of risk like the one computed with the actual conditions would be reached by considering a maximum acceptable critical dilution coefficient r_c of about 5, instead of being equal to 3. Moreover, the implementation of additional IT security solutions, such as a proper firewall for the web application, would impact for instance the scores S9, S12, and S13 of the InfraRisk-CP assessment, providing an estimated significant less probability of successful attack.

The decision about the risk reduction measures to be implemented is dependent on a cost-benefit analysis, thus is highly site-specific. Nevertheless, because of the high potential for risk reduction related to due to the reduction of the probability component, in comparison with the expected implementation cost, the solutions *Information Security Management System* and *Restricted Access to IT system* are suggested for the case study.

By assuming an effective implementation of the measures *Information Security Management System* and *Restricted Access to IT system* to reduce probabilities of a successful attack, the score S5 and S6 would pass from 4 (current situation) to respectively 1 (both measures affect S5) and 2 (Information Security Management System significantly affects S6, but Restricted Access to IT system does not), leading to an estimated f_A equal to 0.063/year, corresponding to a risk equal to 117 m³/year is estimated, i.e., **Low Risk**, considering the selected risk criterion.

A.6. Sample material for chapter: IoT security for water

Topic

This module aims at raising awareness towards the security of IoT devices and more generally of digital solutions using IoT.

The border between IT and OT is getting thinner and thinner as the OT world starts to adopt technologies from the IT world. This comes with a shift of paradigm: the automation pyramid will collapse, and it is vital to be prepared for this event to prevent catastrophic failures from happening.

In this module, the focus is brought on the vulnerabilities of IoT devices and the risks that entail. While most infrastructure have not yet fully integrated IoT solutions into their systems, their simple use can present a risk.

Goals

- Be aware of the risks that come with IoT and the use of digital solutions
- Understand the security testing methodology for IoT devices
- Be aware of the security best practices when developing IoT solutions

Instructor pre-requisites

This topic should be taught by an instructor with competence in cybersecurity and computer network engineering.

Text

Introduction & Background

Internet of Things (IoT) refers to physical objects that run some sort of software and have thus processing capabilities. The IoT umbrella is wide and includes objects ranging from small temperature sensors deployed in the wild to connected fridges, going through medical devices. In the context of this module, IoT will refer to sensors, actuators, and gateways devices, connected to a network (it does not have to be the Internet) used as part of digital solution.

IoT devices often operate on battery and should as such have low power consumption. This comes with the drawback of having limited computing power and as a result, low cryptographic capabilities (amongst others). In addition, IoT devices tend to be mass produced and manufacturers try to keep the cost as low as possible, sometimes saving money on security features (for instance by choosing a microcontroller that does not have hardware support for some cryptographic algorithms). As a direct consequence of these two points, IoT is often associated with poor security.

IoT devices are usually wireless. A wide variety of wireless communication protocols exist, and manufacturers need to choose based on criteria such as the range and the required data rate. Protocols can be split in three families:

- Cellular communication (2G, 3G, 4G, 5G): best range and data rate, but also more expensive and requires more power.
- Short range wireless communication (e.g., BLE,¹⁶ NFC,¹⁷ Zigbee,¹⁸ etc.)
- Low Power Wide Area Network (LPWAN) (e.g., LoRa,¹⁹ NB-IoT,²⁰ Sigfox,²¹ etc.)

Other constraints manufacturers need to take into accounts include scalability, interoperability, cost and security. IoT devices can be connected using different network topologies (bus, mesh, ring, etc.).

Historically, IT and OT systems have almost been two different worlds, using different technologies. However, with the shift to Industry 4.0, more and more domains have started to include IoT-based digital solutions to their ecosystem. This is also valid for critical infrastructures, water included.

For now, though, the data from the digital solutions is usually fully processed outside the utility systems and fed into internal databases and/or numerical models. However, because internal IT security standards might prevent the use of cloud solutions, standalone applications and on-premises servers are sometimes required. Digital solutions usually provide devices, typically sets of specific sensors, mobile applications, or web platforms which, to a large extent, build on top of existing solutions. Sensors themselves are mostly off the shelf, proprietary or open source.

For any sensor implementation it must be distinguished between sensors that are directly connected to the SCADA system to monitor and regulate operational conditions, or independent installations in online or offline mode with data transfer into the databases and reporting systems connecting to the office world. Data transformation and processing is accordingly done either fully automated within the SCADA system or manually to semi-automated when reading, recording, and processing the office world data. The OT and IT systems being usually not directly connected, the data is pushed from the operational systems to central databases, where they are stored and accessed for decision making.

Digital Solutions included in water infrastructures so far are built using a wide range of technologies, making it difficult to generalize anything. However, when discussing with technology providers and water utilities, it is possible to classify solutions in three subsets, based on their level of integration with the water utilities:

- **Standalone solutions** These are solutions which are not interacting with any sensors or utilities. They can be Web or Mobile Applications, publicly available or requiring authentication.
- **Solution with external sensors** These are the solutions that are gathering data from sensors “in the wild”, using some long-range wireless technology (mobile networks, LoRa or Sigfox for instance) or manually gathered using most likely shorter-range wireless protocols, such as Bluetooth (Low Energy).

¹⁶ <https://www.bluetooth.com/learn-about-bluetooth/tech-overview/>

¹⁷ <https://nfc-forum.org/learn/nfc-technology/>

¹⁸ <https://csa-iot.org/all-solutions/zigbee/>

¹⁹ <https://lora-alliance.org/>

²⁰ <https://www.gsma.com/iot/narrow-band-internet-of-things-nb-iot/>

²¹ <https://www.sigfox.com/en/what-sigfox/technology>

- **Solution with *internal* sensors** These are the solutions that are gathering data from sensors that are placed in the water utilities (but not necessarily connected to their systems).

Standalone solutions that do not rely on IoT devices are out of scope for this module. This leaves solutions with *external* sensors, and solutions with *internal* sensors. These solutions still rely on a whole stack to provide value to the water utilities (i.e., a sensor but also a companion app for instance). Even if a digital solution is not directly connected to the water utilities, there will still be interactions, especially by operators accessing the application to take decisions. As such two major risks for water utilities can be identified when integrating/using digital solutions:

- **Supply-chain attacks** As already mentioned, water utilities being critical infrastructures, they must comply with strict regulations. It is safe to assume that they undergo regular audits and are supposed to be a difficult target to find an entry point to for attackers. This is the case for many industries and companies. To counter this, attackers might identify actors evolving around the main target and attack these companies instead, to later on leverage the trust in these companies as a mean to attack the real target. This is called a supply chain attack and has been used successfully by attackers in the past. The most damaging one in the past years is the SolarWinds Supply Chain Attack discovered in December 2020. The Network Management System used by hundreds of thousands of companies, was shipped with a malware, successfully hitting many high value targets such as the US Federal government [29]. In January 2021 for instance, vulnerabilities in Microsoft Exchange server were used to compromised hundreds of thousands of servers all around the world, including the European Banking Authority and the Norwegian Parliament [30, 31]. More recently, there has been several cases of malicious node packages being uploaded to npm²² as an attempt to target companies such as Azure, Uber or Airbnb [32].
- **Being led to take wrong operational decisions** Most of the digital solutions, while not interacting with the water utilities, provide crucial information to the operators of a water utility and are used as a decision support tool. The interoperable decision support system and real time control algorithms for stormwater management for instance, can be used by operators to predict the best maintenance window, thus optimising the process. If the information is erroneous though, it can lead to a release of untreated water in the environment as a direct result of the wrong planning. Similarly, if the early warning system for bathing water quality reports an incorrect value so that bathing is authorized despite the quality being below the threshold, this can have disastrous consequences (both from a public health perspective, but also from a public relation one).

Digital solutions vary a lot when it comes to the technologies used and to the services they provide, but it is possible to derive a high-level diagram of their architecture. Figure 19 presents such a diagram. Most solutions are only composed of a subset of the components presented here. The main components include:

- **Data sources** Digital solutions usually rely on external data, which is then analysed and/or transformed to provide added value. This data can take various shapes: some solutions collect

²² npm is a package manager for the JavaScript programming language

environmental data using IoT sensors deployed in the wild (water sources, sewer network, etc.), others use data from 3rd party services (weather or terrain information for instance) or even drones. The applications developed by a solution can themselves be considered as a data source; for instance, when collecting data from an off-line sensor using Bluetooth.

- **Solution infrastructure** Most solutions rely on a backend infrastructure to operate their service. Infrastructure here refers to anything that supports services run by a digital solution and can consist of on-premises servers, cloud ones but also 3rd party services used as part of the data collection (Sigfox's network for instance). Network providers used for data collection are here considered as part of the solution infrastructure, contrary to the other external services described in the next point.
- **3rd party services** These are the 3rd party services used by a solution to provide their own service, such as services providing SMS or email sending capabilities.
- **Solution's services** Solutions provide a service to water utilities/their users. This can be for instance an alert if the level of the E. Coli bacteria is too high in a water basin. A service can be exposed to the users in various forms, such as a web or mobile application, but also simply via an Application Programming Interface (API).
- **Users of the solution's services**
 - **Regular users:** Users of the service are for instance operators in a water utility in need to take operational decisions based on the information they receive from the digital solution's service. A good example could be operators visualising on its application that the level of E. Coli bacteria is higher than a given threshold in a water basin and deciding to forbid swimming there.
 - **Machines:** If the service is exposed via an API, it might be used by another solution to develop something novel, or directly by a water utility to integrate it within their own system.

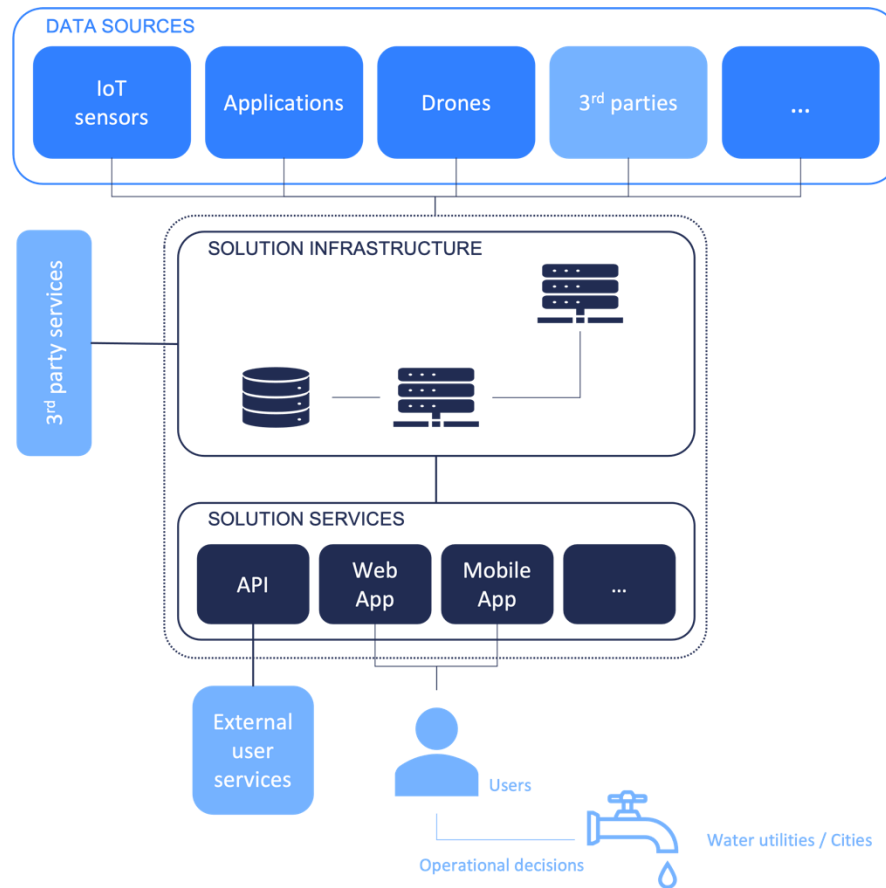


Figure 19 Generic Architecture Diagram for a Digital Solution (in DWC)

Threat Landscape for Digital Solutions

Motivations for attackers to attack digital solutions exist, independently of their level of integration with water utilities, and this justifies the need to ensure they are secure as well [33, 34].

As presented in the previous section, digital solutions can be complex and might interact with several external actors to collect data, access services or to simply provide their own service to their users. Attackers thus get a wide choice of attack vectors when targeting a digital solution: IoT devices, applications, 3rd party services, etc.

Using the Risk Identification Database (RIDB) which gathers the generic risk events associated with the implementation of the digital solution of DWC by the cities, one can derive a classification of the attacks to the digital solutions in DWC. This classification is presented in Figure 20, and groups the attacks in 6 different classes which sum up the different types of attacks that can relate to a digital solution:

- **Attacks on IoT sensors:** As already explained above, IoT sensors are for many solutions a keystone as they constitute the source of data. Sensors are particularly vulnerable since they are often deployed in the wild (i.e., accessible by almost anyone), and are difficult and costly

to secure (and thus have historically had poor security). While the data of the sensors themselves might not be very valuable to an attacker, being able to manipulate this data to trigger incorrect outputs by the services could have severe consequences. In addition, IoT sensors have in the past (and keep being so) already been compromised at scale to be integrated in botnets [16].

- **Attacks on infrastructure:** attacking the infrastructure is a way for the attackers to either disrupt the services (by for instance launching Denial of Services (DoS) attacks on it) or to gain unauthorized access to resources. This could also be a way for attackers to later use their access to a service to attack a water utility using and trusting this service.
- **Attacks on ML/AI:** compared to the other types of attack in this classification, attacks against Machine Learning (ML) and Artificial Intelligence (AI) are less known and appears a bit as newcomers. Attackers can for instance use specially crafted inputs to mislead the algorithms. Famous examples of such attacks include for instance cars being tricked into speeding by placing tape on speed signs.²³ Other attacks consist of an attacker feeding incorrect data to the classifier, polluting the model in such way that its own data is later classified as good data, or on the contrary so that good data is inf act classified as incorrect. Finally, models have an intrinsic value, and an attacker might want to steal them.
- **Attacks on applications:** Applications (Web, Mobile, API, etc.) are usually exposed and if compromised, can lead to data leak, unauthorized access to resources and actions or allow for data manipulation and denial of service.
- **Human errors/failures:** while not being an attack per se, human error can lead to the same consequences. If a user is given access to data or actions, he/she should not have access to, he/she could misuse it (intentionally or not) and effectively create a situation like an attack (for instance, a user could be given access to an alert system and trigger an alarm, leading operators to take decisions based on misleading data).
- **Social engineering:** Like human errors, a user could be tricked by a malicious person into performing harmful actions, potentially leading to dangerous consequences as well.

²³ <https://www.mcafee.com/blogs/other-blogs/mcafee-labs/model-hacking-adas-to-pave-safer-roads-for-autonomous-vehicles/>

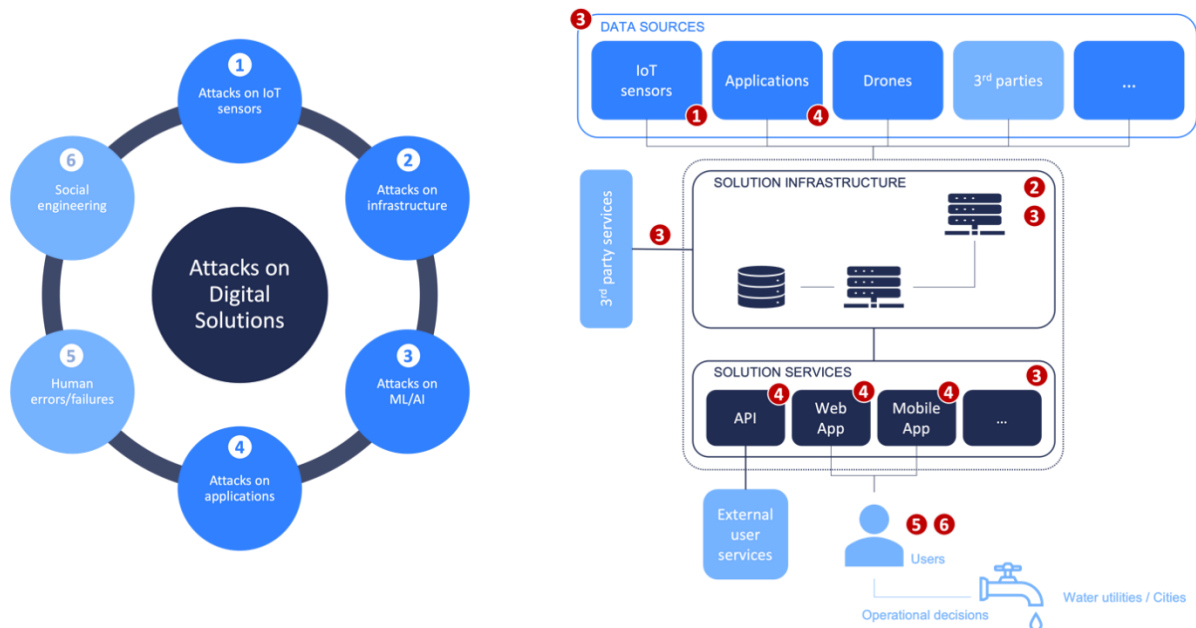


Figure 20: Classification of the attacks against Digital Solutions (in DWC)

When securing a digital solution, it is important to take into consideration who and what one is defending against. Indeed, defending against a high-school student, running automated tools he found on the Internet, is not the same as defending against a state-sponsored threat actor that has unlimited resources. As presented by Weingart [35], we can classify attackers into three different classes based on their capabilities. While his classification is intended for physical security of embedded devices, it can be slightly adjusted the context of this course:

Class I A clever outsider, who has limited knowledge about the system and a low budget and equipment. This could be a curious attacker that is targeting the system mostly for prestige and as a hobby, but also a *script kiddie*,²⁴ dumbly following scripts and tutorials found on the Internet.

Class II A knowledgeable insider, who has advanced knowledge and/or specialized education and experience in the area. This category has access to sophisticated tools. Typically, this class corresponds to researchers.

Class III A funded organization categorized by its high budget and its ability to recruit class II attackers to attack the system. This corresponds to organized crime or to a government.

Attackers having incentives to target digital solutions to disrupt or gain access to water utilities, it is thus expected digital solutions should consider attackers from all three classes. However, for many

²⁴ A script kiddie is person who uses existing computer scripts or codes to hack into computers, lacking the expertise to write their own.

solutions, defending effectively against state-sponsored actors is simply not feasible, as it would drastically increase the cost of their product, making it too expensive for any water utility to adopt. This is especially true for IoT based solutions, which rely on low-cost environmental sensors to collect data. Securing these devices while keeping the cost low is a challenge given today's state of the art in the area: the devices being deployed in the wild most of the time, they can be accessed by a malicious person and analysed not only from a software, but also from a hardware perspective. Microsoft's third immutable law of security states that *"if a bad guy has unrestricted physical access to your computer, it's not your computer anymore"* [36] and this holds even more true in the context of IoT devices. This, however, does not mean the solutions should give up on security, as they can still protect against class I and class II attackers, by for instance tackling the low hanging fruits in their product.

To secure a product against all classes of attackers, a change of paradigm is required: one must work with the assumption that parts of the solution will be broken/accessed by attackers (typically an IoT sensor) and ensure that the impact of this breach has no operational or financial consequences. There is no "one size fits all" scenario, and solutions must assess on their own where they stand, and how much effort is needed to secure their product. Indeed, a solution owner might be concerned by the Intellectual Property (IP) an attacker could get his hands on if he were to compromise an IoT device (models, algorithms, etc.) and thus choose to invest in more hardware security than for another solution that only measures environmental data to send them back to a backend infrastructure for processing.

Another way to think about the problem is through cost: attackers, no matter which class they belong to, will go for the easiest and cheapest path that has the most impact. As such, ensuring the low hanging fruits are tackled will increase the cost and difficulty of an attack. Reducing the impact (by for instance ensuring proper segmentation) also contributes to attackers looking elsewhere.

Security Testing Methodology for IoT Devices

For the digital solutions' owners to be able to weigh the cost and benefit of their different options, they first need to be aware of the risk magnitude. While security must be considered as part of the design process (*security by design*), having a way to ensure that there are no major security flaws in the product is extremely valuable. Indeed, even if the solution has been designed with security in mind, programming errors or misconfigurations are common and can cancel out those efforts. In this section we present different approaches to (practical) security testing of a solution.

Penetration testing

Penetration testing is defined by ENISA as *"the assessment of the security of a system against different types of attacks performed by an authorised security expert. The tester attempts to identify and exploit the system's vulnerabilities. The difference between a penetration test and an actual attack is that the former is done by a tester who has permission to assess the security of the system and expose its security weaknesses. In addition, the tester is given certain boundaries to operate and perform this task."* [37]

Penetration testing is usually divided into three subcategories, based on the level of knowledge the security expert is provided with at the start of the assessment: white, grey, and black box testing.

White box testing

In a white box testing configuration, the security expert performing the assessment is given full knowledge of the tested system (code, architecture documents and other useful information). The objectives of a white box testing are to simulate an insider’s attack (for instance a previous employee who has had access to the code or even developed part of it). The advantage of a white box testing is that it might uncover bugs and vulnerabilities that can be difficult to identify by the other tests within a short timeframe. This benefit can also be seen as a disadvantage, as some of the uncovered vulnerabilities might not be “real vulnerabilities”, in the sense that they cannot be exploited by an attacker (one can for instance imagine a function being flawed and leading to a buffer overflow when called with the wrong parameters; but if it is only called with constant values by the programmer, a malicious attacker will not be able to trigger it).

Grey box testing

Grey box testing is a mid-step between white box testing and black box testing: the security expert is given some information regarding the system targeted. This can be for instance indications on the technology used in the backend, available services (supposed to be non-accessible for instance) or some of the algorithms used. The pro of a grey box testing is that it combines some of the “real-world” advantages of the black box methodology (see below), while being more effective (the security expert does not lose time testing for things he could have known directly by looking at the documentation).

Black box testing

In a black box testing scenario, the security expert has very little or no previous knowledge of the targeted system or device. At a high level, the methodology consists of sending inputs to the system, the “black box”, and to analyse the obtained outputs to deduce the internals of the target. Having made some guesses, the attacker can adjust her inputs to confirm her thoughts or to exploit the target. This is presented in Figure 21.



Figure 21: High level diagram of the Black Box Testing Methodology

The black box testing methodology has several advantages over the white and grey ones. Indeed, its primary objective is to test a system under real conditions, to emulate a real attack scenario. This means that such a test might catch errors made during the deployment of the system such as default passwords, misconfigurations in general or even the lack of security trainings of operators (weak passwords). This methodology also presents a low false positive ratio as the security expert can assess the risks associated with a vulnerability directly, i.e., if the vulnerability can be exploited or not. While black box testing can miss some vulnerabilities, it should likely be the first test performed. It is an excellent way to assess how a system stands against attacks and to get an idea of the path an attacker

would take to compromise the solution, and thus gives indications on how to tackle those potential low hanging fruits. It can be later be completed by a deeper assessment following a grey or white box approach.

Red team exercise

In a penetration test, there is no element of surprise, and the scope is limited to the system being assessed. A red team exercise goes one step further and includes physical penetration and social engineering: the objective of the attackers is to compromise a target by all possible means without being detected. This allows the company/organisation being tested to not only detect vulnerabilities in their system (like in a penetration test), but also to test their detection and response capabilities. In the context of critical infrastructures, here water infrastructures, red team exercises would be more targeted towards mature organisations such as water utilities.

Best practices for IoT-based Solution Development

This section gathers best practices and recommendations to digital solutions' providers for a secure integration with the water utilities. Solutions considered here are the ones from the DWC project. Given that they have little to no integration with the SCADA system (see D4.1 [46]), the recommendations come down to secure development of a digital solution.

We provide general recommendations for security that apply all along the solution lifecycle (from design and implementation to operation). As mentioned previously, digital solutions in DWC are diverse when it comes to the service they provide, the technologies they use and how they integrate with the utilities. As such, it not possible to cover all security measures in an exhaustive way.

General recommendations

Back in 2000, Bruce Schneier wrote *"Security is a process, not a product. Products provide some protection, but the only way to effectively do business in an insecure world is to put processes in place that recognize the inherent insecurity in the products. The trick is to reduce your risk of exposure regardless of the products or patches."* [38] This still stands strong, if not stronger, by today's standards.

The following recommendations aim to help build a security process when developing a digital solution. They have been grouped in three different categories covering the full lifecycle of a solution: its design, implementation, and operation. Those recommendations are generic and apply to any solution.

Designing

Know your enemy

The first step in designing a secure solution is to get an understanding of who and what to defend against. Companies might not always be aware that they can be targets, but as already mentioned, they can sometimes simply be used as a means to reach the primary target (supply chain attacks). As

such, companies must have an idea of the different profiles of the attackers, from a skilled hobbyist to a state sponsored hacker group or even an employee made unhappy by being denied a raise, or recently fired.

While it is important to know the profiles of the attackers, it is maybe even more important to understand how they operate or, in “technical terms”, what are their tactics, techniques and procedures (TTP) [39]. It is advised to keep up to date with threat intelligence provided by security companies such as Mnemonic or Trend Micro.

Overall, it is advised to acquire an adversarial mindset when developing a solution. Having an idea of what is possible, or simply what are the different attacks might not be an easy task when one has no security background. Frameworks like the Mitre ATT&CK²⁵, a knowledge base of adversary tactics and techniques based on real world-observations, can help getting an overview of how attackers operate in the wild.

Define your security requirements

When building a new product, one of the first stages is to define a set of requirements for the new system. One usually focuses on *functional requirements* which define the different features of the products, but non-functional requirements such as security and reliability requirements also come into play. Some of those features might come from external sources, such as customers or regulations, and might at first seem in conflict with some of the functional requirements. One might then have to make trade-offs. It is often tempting to postpone dealing with security to instead focus on the functional requirements, but security is not optional for services in production, and a lack of it can lead to loss of business or worse, bankruptcy of the company [40]. Ensuring security by design might have a slight over-cost initially, but it pays out eventually.

Design for least privilege

The *principle of least privilege* states that “users and programs should only have the necessary privileges to complete their tasks.” [41] This limits the attack surface for attacks (internal attacks or compromise) and mistakes. In practice, this principle can be applied directly in the design by making each component/program do one thing well, and one thing only. This allows the application of the principle, by granting the necessary permissions to use a specific program (for instance an API endpoint). In addition to this, it is important to have good audit capabilities of the system to detect any malicious use of it, or simply to understand what happened in case of an incident, should an incident occur.

Design for a changing landscape

A product is usually designed with longevity in mind, and as such, the company is prone to see several changes in the threats to their product, but also in applicable regulations and expectations from their clients: those expectations are much likely to rise than to go down. In addition, a company/product might grow, and with this growth will likely come more security requirements. Past events have shown

²⁵ <https://attack.mitre.org/>

that companies failing to adapt, or having a poor initial design are likely to fail: some have even been forced to go bankrupt after a successful cyberattack on their system. This is for instance the case for “The Heritage Company” who had to shut down back in December 2019 after a ransomware attack, they did not manage to recover from [42]. This highlights the need to design a product with these considerations in mind. In practice, this translates to building an infrastructure which makes changes easier: for instance, using containers helps keeping the application and its dependencies in a single package decoupled from the underlying OS, easing the update process (one can simply patch/update the image in the container registry, effectively helping to patch as part of the code deployment process). Another potential recommendation is to use microservices when suitable: splitting services in smaller units helps with scalability, maintenance, and security patching.

Design for resilience

Companies should strive for their solution to be resilient i.e., prepared to handle the unexpected. This can be a cyberattack or simply a legit increase of traffic for instance. Some strategies and mechanisms can help with this. Defence in depth is one of them. It consists in applying several defence mechanisms in a row, thus forcing an attacker to defeat all of them before reaching his goal. In the case of a web application, this could be for instance first a web application firewall, then access control on the API endpoints of the application, and perhaps containerization of the service. An attacker would then have to bypass all measures to get access to the host machine. In addition to defence in depth, it is recommended to have automated response mechanisms in place. This could for example be load shedding (a service replies with an error code if overloaded) or client throttling (increase of the response time) in the case of web service. Finally, even in the case of a compromised service, it is important to have mechanisms to reduce the impact, the most efficient being perhaps segmentation and strict application of the least privilege principle already mentioned above.

Additional considerations

While building security in one’s product is important, one must sometime be pragmatic and consider the solution developed in its global context. Indeed, too much security can sometime have the opposite effect of what is expected. The best example for this is probably a password policy being too complex, which leads to users not able to remember their passwords, and them ending up writing those passwords down on a sticky note fixed on their screen. This happened for example to the French channel TV5 Monde, whose staff accidentally showed their passwords during an interview [43]. Similarly, security might impact the fluidity of operations, especially in a crisis. Facebook’s outage in October 2021 is good proof of it: a configuration error on a BGP router disconnected Facebook’s network from the Internet, and due to strict security policies, engineers faced difficulties to access the data center, and later on to access the hardware itself to solve the issue.

Implementing

Ensure security by design

As already mentioned in the design section, security must be thought at the design stage, but it should also be the default option chosen when developing the solution. As an example, one can think about a solution exposing a service over HTTPS. While there could be reasons to allow the user to disable

HTTPS and use HTTP, the default option should be to use HTTPS. Similarly, one must ensure that the solution will fail securely, i.e., if a piece of code fails, the software fallbacks to a secure state. The following pseudo-code taken from OWASP²⁶ exposes the problem clearly:

```
isAdmin = true;
try {
  codeWhichMayFail();
  isAdmin = isUserInRole("Administrator");
}
catch (Exception ex) {
  log.write(ex.toString());
}
```

If either the `codeWhichMayFail` or `isUserInRole` function fails, the user will be administrator by default.

Use frameworks

Whenever possible, one should not reinvent the wheel, but rather rely on existing frameworks. Writing applications that maintain security properties quickly become challenging as the project grows. Frameworks often provide all the necessary boilerplate to deal with common tasks regarding security (user authentication and authorization, logging, XSS and CSRF protections, etc.). Not having to re-create those basic blocks for every new solution helps reducing costs, but also helps with maintenance. Examples of such frameworks for web applications include Django, Laravel or Spring Boot. This recommendation also applies for embedded systems where embedded OS / Real Time Operating Systems provide developers with the means to implement things securely (ZephyrOS²⁷ and FreeRTOS²⁸ are such examples).

Be aware of common security pitfalls

During the implementation phase, it is recommended for engineers and developers to have common security pitfalls in mind. Initiatives such as the OWASP Top 10²⁹ or the SANS Top 25³⁰ can provide food for thought for engineers to build upon. Checklists³¹ provided within DWC also aim to help building awareness regarding common security issues, not only in web applications but in the different areas highlighted in the previous section.

Test your solution

Solutions should test their solution as much as possible and using different approaches. To err is human, and even the most careful engineers make mistakes or forget about edge cases. In practice,

²⁶ https://wiki.owasp.org/index.php/Security_by_Design_Principles#Security_architecture

²⁷ <https://www.zephyrproject.org/>

²⁸ <https://www.freertos.org/>

²⁹ <https://owasp.org/www-project-top-ten/>

³⁰ <https://www.sans.org/top25-software-errors/>

³¹ To see all the checklists provided as part of DWC, see D4.3

this means that there are many ways for an untested solution to fail in the real world. Some of the recommended testing techniques include:

- *Unit testing*: breaking the software in a small testable unit that have no external dependencies. At that stage, external components such as databases are mocked. Most languages have frameworks to help with unit testing.
- *Integration tests*: those tests take place one level above the unit tests and use real implementation of the databases and other components. These tests are about ensuring components integrate properly and securely together and that the system as whole behaves as expected.
- *Static code analysis*: as the name indicates, static code analysers inspect the source code of a program to detect potential errors or vulnerabilities. They can help catch issues before the code is built and deployed.
- *Dynamic code analysis*: this approach analyses the developed software in a dynamic manner to profile the performance, evaluate code coverage or security correctness. Dynamic analysis can identify several errors as well, such as race conditions, initialized memory being read, memory leaks or out-of-bounds memory accesses to quote a few.
- *Fuzzing*: fuzzing complements the above-mentioned techniques and consists in generating large amounts of inputs to the components being tested and analysing how it reacts to those inputs. It is particularly useful to detect bugs in parsers or protocol implementations. Fuzzing is even included by default in some languages such as Golang.³²

Ensure secure deployment

It is best practice for new code added to the project to be reviewed before being pushed to production (or even checked in). Enforcing code review for every piece of code also helps protecting against malicious employees. Code review must then be mandatory, and one should not be able to opt out. This not only helps catching potential security flaws but also “regular” bugs and errors.

When deploying code, one should rely on automation as much as possible: automation removes the human from the loop, thus preventing mistakes and providing a consistent, repeatable process for building, testing and deploying software. [44]

Operating

Perform Penetration Testing and Red Team Exercises

It is important for solutions to have their solutions tested by performing both penetration testing and red team exercises. As already explained in the methodology section, those two methodologies are complementary and provide a good overview of the “real-life” security of the product. One could also imagine that water utilities might be interested in using Red Team Exercises to get an assessment not only of their technical security level, but also of their operational preparedness.

³² <https://go.dev/doc/fuzz/>

Prepare for vulnerabilities

New vulnerabilities are disclosed every day, either in a software used directly by a solution (OS, container technology, etc.) or in a library used to develop the solution. Vulnerabilities can also be discovered in the solution itself. As such, it is important for a solution to be prepared to handle vulnerabilities (either the ones impacting the solution, or the ones within the solution). This means having some sort of monitoring in place to be aware of the new vulnerabilities, but also a plan to patch, disseminate and communicate about vulnerabilities found in the solution with their users.

Have a crisis management plan

While all the measures described previously in this section aim at preventing security breaches, one must plan for the worst, and as such for a security breach. If such an incident should happen, one must have a plan to deal with it in order to keep the system up and running in a secure and reliable manner if possible. This is a huge topic in itself, and we invite the solutions' owners to refer to the Google SRE book linked below.

While this whole section provided a wide range of information regarding security in the design, implementation and operation phases of digital solutions, its objective was to give an overview of the different activities to perform and have in mind, and as such, barely scratched the surface. The following resources were used to build this list and are recommended to dig deeper on the topic:

- OWASP— Security by Design Principles³³
- OWASP – Vulnerability Disclosure Cheat Sheet³⁴
- Google SRE— Building secure and reliable systems³⁵
- OWASP SAMM v2³⁶
- Security Engineering Book³⁷
- ISO/IEC 27034³⁸

Common attacks and security vulnerabilities on sensors/IoT

Attacks against IoT represents a fair part of the potential risks identified in the Risk Identification Database of DWC. IoT solutions in DWC are similar to other industries when it comes to the technologies used and to their design. As such, we find the same issues, and can also provide the same advice. We have prepared a checklist for IoT Security³⁹. The checklist questions are divided in a generic section and in 4 main areas that constitutes a high-level division of most IoT solutions:

- **Hardware:** the device itself (PCB design, debug ports, etc.).
- **Software:** the code running on the firmware (not the code running on the infrastructure).

³³ https://wiki.owasp.org/index.php/Security_by_Design_Principles

³⁴ https://cheatsheetseries.owasp.org/cheatsheets/Vulnerability_Disclosure_Cheat_Sheet.html

³⁵ <https://sre.google/books/building-secure-reliable-systems/>

³⁶ <https://owaspsamm.org>

³⁷ <https://www.cl.cam.ac.uk/~rja14/book.html>

³⁸ <https://www.iso.org/standard/44378.html>

³⁹ The latest version can be downloaded at <https://www.sintef.no/en/projects/2022/ragnarok/outcomes>

- **Communication:** how is the data acquired and transmitted.
- **Infrastructure:** the backend infrastructure, servers, application, etc.

This organisation of the checklist highlights another important aspect of IoT solutions: they require a broad range of expertise (from hardware to backend going through embedded system development) and as a result teams are usually only experts on their specific topic and might work with assumptions on what other teams do, especially when it comes to security. It is thus vital for solutions that the teams have a good understanding of what level of security each team can provide and not to work with assumptions.

In addition to our checklist, existing guidelines such as the "[Baseline Security Recommendations for IoT in the context of Critical Information Infrastructure](#)" [45] from ENISA can be used to get a better understanding of the risk linked to IoT and of the countermeasures that can be applied.

A.7. Sample material for chapter: Scenario-based training exercise

Topic

This course module aims at creating cyber security awareness through scenarios-based training.

The training will make use of a role Game for Cyber Security Training developed by the STOP-IT project and further extended and tested in the DWC project. The game is called TORC.

More than a teaching module, this chapter provides all relevant information to guide the creation of a training session through the use of the TORC game. The material provided include a general overview about TORC, instructions on how to play, with characterization of the different roles, and finally two scenarios are also provided as examples for trainees interested to create their own scenarios following the proposed template. The first scenario was developed in the STOP-IT project, while the second is derived from the DWC activities documented in D4.3 [27].

This chapter could be used by any (water) organization aiming to complement training activities of employees to increase organizational preparedness on cybersecurity and for encouraging behavioral change. It can also be seen as a training exercise to be organized as the last module of the training scheme here proposed (e.g., what learnt in previous modules can be applied while playing with TORC).

The dissemination video developed in STOP-IT can support the understanding of the scope of the game, as well as be used to introduce the training exercise: <https://youtu.be/qwGuLfdYzJw>

Goals

- Facilitate a case-based learning approach
- Provide instructions to trainees on performing scenario-based training and in turn:
 - Increase cybersecurity awareness by playing
 - Change the player's long-term security behavior

Instructor pre-requisites

This topic should be taught/facilitated by an instructor with competence in water network engineering (depending on the specific case chosen for the exercise).

Text

Overview of TORC

TORC is designed to facilitate organizations and teams that seek to reveal, understand, articulate, demonstrate and/or develop their inherent repertoire of risk management and/or resilient performance in face of unexpected deviations, disturbances and shocks. The training outcomes and experiences are captured in a way that prepares them to be used as raw material of technological, human, organizational and managerial priorities and resources that are needed to transform the experience from the training exercise into effective resilience capabilities under a more formal

managerial supervision. This approach enables water utilities to develop their human skills while enjoying the benefits of a gaming approach and utilizing competitiveness between the trainees. The foresaid will contribute to the significance of the training.

TORC is available as a digital game, thus players can play remotely from each other.

The TORC game is available at: <https://digital.torc.no/>

The practical instructions on how to play the digital TORC are provided at the following link: <https://documentation.torc.no/>

A TORC game session is suitable for training groups up to 15 people. A substantial improvement can be achieved for autonomous training of a single trainee group with a common challenge and joint field of practice. However, individual training groups and results can be combined interactively across organizational levels and domains and dispersed in time to facilitate a broader strategic objective of resilience training and development beyond the confines of the individual training group. Moreover, TORC training groups may be composed homogeneously or heterogeneously depending on the overall (resilience) objective and needs of the trainees' organization or training sponsor.

The simplicity of the TORC approach and gaming material per se is somewhat counterweighed by the need to prepare detailed training material for specific training contexts, e.g., specifications of the operational situations subject to potential disturbance, and the specific disturbances that emulates the "surprise" for trainees. The development and coordination of an overall strategic objective for several individual training activities may also seem overwhelming at first glance. However, if an incremental rather than a "grand design" approach is chosen, the efforts as well as the rewards may instead develop more organically over time.

Hence, a substantial benefit may also be derived from such a process of preparation, e.g., a clearer understanding of operational vulnerabilities as well as of mitigation options that may constitute the grounds for effective policies, strategies and objectives for resilience development, training objectives included.

Playing the TORC game

Before starting a session, the TORC manager interface (<https://my.torc.no>) allows to customize the steps played throughout the game and the cards which can be selected by the players to take actions.

In the following, a selection of relevant steps for the so-called "risk mode" is adopted and discussed in detail. However, other steps from the TORC manager can be adopted for other game session when different purposes are relevant (e.g., selection of steps related to the so-called "resilience mode").

The selected TORC steps are illustrated in Figure 22. Upon launching the game, activities within "Framing and setting the scene" step take place. The facilitator assigns a role to each participant which may or may not be his/her actual daily role or affiliation. The facilitator introduces a predefined scenario which will be played out in the training session.

As the first disturbance (stressor) is introduced by the facilitator, the group embarks on process comprising a set of cognitive aspects: awareness building, sense-making, and decision making. This step is arguably the most important step in the game starting from building a common understanding how a detrimental scenario may develop, realizing the impact in the context of the service provided by their organization, assessing the possible risk reduction measures and their effectiveness to overcome the challenges, anticipating the requirements for implementing the solution, and taking the

decision on which alternative(s) to implement. An additional orchestrated scenario can thus be played out by introducing new stressor(s).

The game facilitator decides time constraints and circulation of roles within the trainee group. When the game is concluded, the facilitator and the players summarize the findings of the game. Finally, the facilitator and the players can drive conclusions of the resulting strategies applied during the game and assess if the decisions made are aimed towards detection, prevention or mitigation actions.

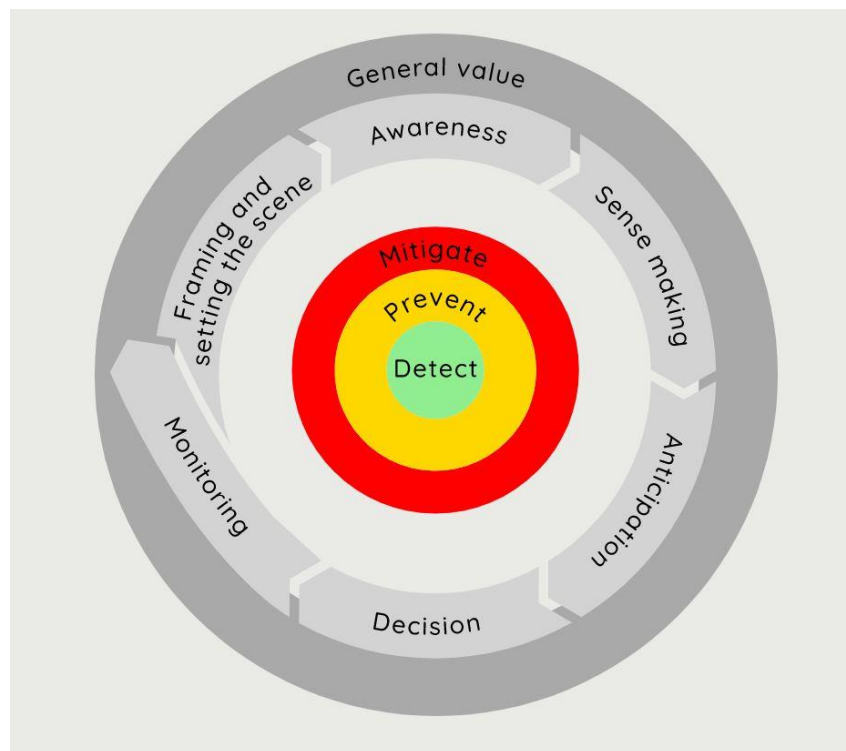


Figure 22: Key features of playing the TORC game

The objectives of each step can be briefly described as follows:

Launching the game:

- 1- Framing and setting the scene: the objective of this step is to provide a customized scenario to players with defined roles including local conditions and the stressors.

Reflection in action process:

- 2- Awareness: This step assesses, as holistically as possible, what might happen, how the scenario and the disturbance might develop into an undesired event.
- 3- Sense making: the assessment of scenario's impact and probabilities on the service or the infrastructure through proper tools selected by the players (selecting proper action cards).
- 4- Anticipation: this step provides a set of possible risk reduction measures, described as action cards, to detect/prevent the stressor and/or to mitigate the consequences of the scenario by evaluating their effectiveness.
- 5- Feasibility: this step is optional and studies the technical and organizational feasibility requirements to implement the possible alternatives.

- 6- Decision: within this step, one or a set of alternatives are chosen to be implemented, having the possibility of tagging the final decision as a solution aimed at detecting the risk event, preventing the risk event, or mitigating the consequences of the risk event.

Reflection and discussion after action process and evaluation of the game result:

- 7- Monitoring, and general value: the final phase summarizes the findings also in a long term perspective.

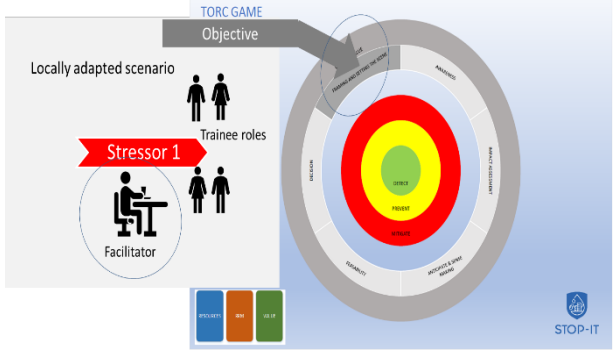
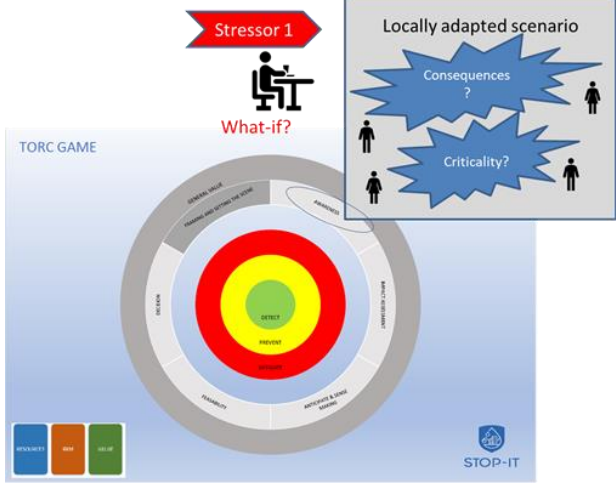
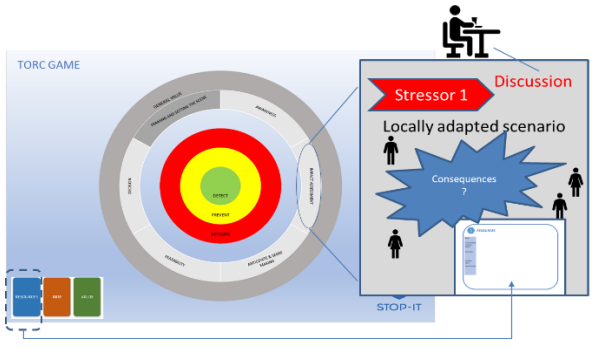
The typical single TORC training session in "risk mode" unfolds as follows:

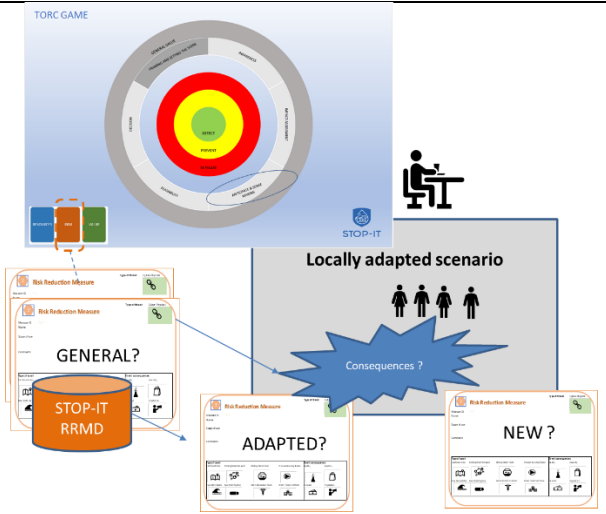
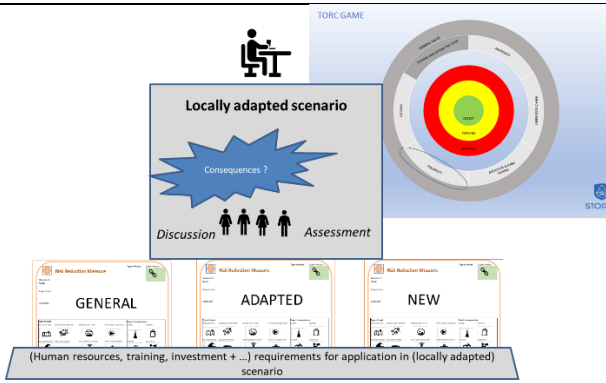
- A training situation is defined as a scenario being on the edge of normalcy. That is, it is conceivable to introduce a disturbance (stressor) that renders the normal preparations and ways of working insufficient to deal with the possible unwanted consequences.
- The training group is composed in a manner to ensure that each of the members are able to recognize the essence of the scenario, identify possible unwanted consequences, and be associated with some level of practice that can influence the situation (after a disturbance is launched).
- A disturbance (stressor) is launched by the facilitator. The group is given a time limit to go through all steps of the game board. One individual trainee is assigned to be the head of the trainee group (the decision maker).
- The game starts at the "awareness" field. Here, the group is expected to identify additional information to understand as much as possible of the disturbance and identify ways of gathering such information.
- The next step is "sense making". Here, the group is expected to elaborate and describe the potential (undesired) impact of the disturbance, how it may evolve, how likely it can occur, and identify action cards that could be used to better assess the risk.
- The next step is "anticipation". Here, the group is expected to elaborate and describe various potential alternatives for mitigating the risk generated by the disturbance. The group selects existing action cards and/ or create additional new cards corresponding to mitigation actions.
- The optional next step is "feasibility". Here, the group assesses the technical and organizational feasibility requirements to implement the possible alternatives.
- The next step is "decision". Here, the group is expected to select one or more of the alternative actions. If the group does not agree, the head of the group must decide.
- If a cascading training is wanted, the facilitator issues a new disturbance, the role of head of the training group is shifted, and the whole process is repeated. The new disturbance might be predefined, determined by the facilitator due to the actual circumstance, or derived from the actions of the trainees themselves (e.g., activating an issue that the trainees has identified for instance through the preceding "feasibility" step.

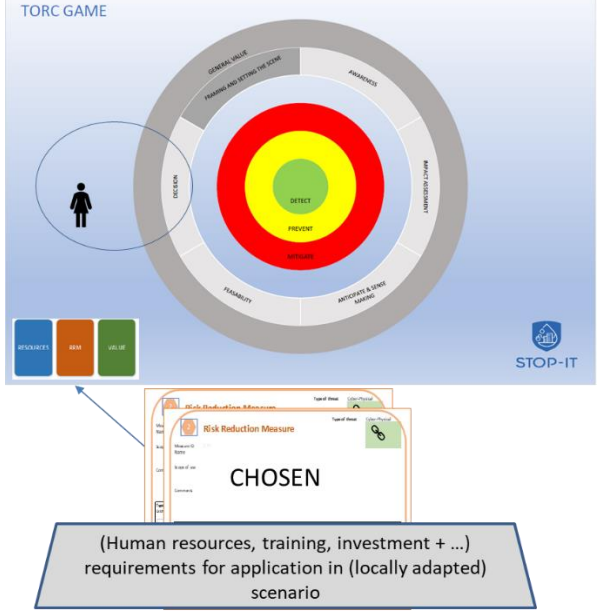
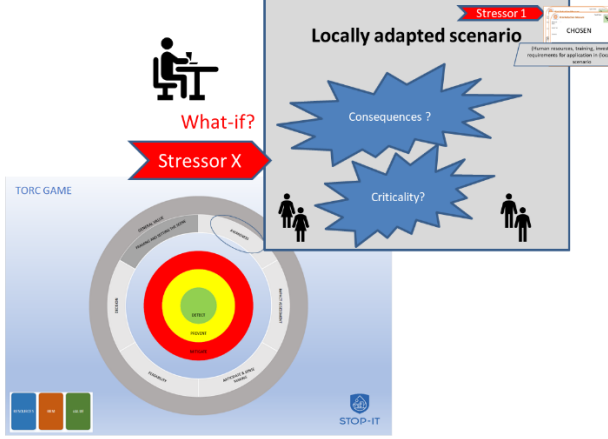
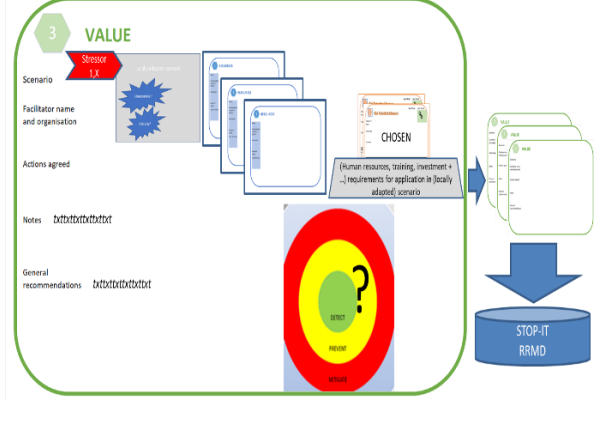
For each of the steps, resources used, skills and competences deemed necessary, and plans for coordinated actions taken are described on a final report and put on the game log contributing to the inventory.

When the training session ends, an evaluation of the game results is conducted by reconstructing the training activity and reflecting on the general value of the session. Table 5 describes the specific actions corresponding to each step of the game and the role of different players.

Table 5: TORC quick-start guide (illustrations derived from STOP-IT project [28])

Game phase and activity	Illustration
<p>1) Framing and setting the scene</p> <ul style="list-style-type: none"> The facilitator prepares prior to the game a customized setting (stressor-scenario) adapted to the context of the concerned utility playing the game. The facilitator points out the roles of each player, then, he introduces the infrastructure's context. The facilitator then introduces a predefined stressor to the system as a part of the predefined scenario. 	
<p>2) Awareness</p> <ul style="list-style-type: none"> The trainees discuss the actual criticality of the stressor and what are the possible scenario's consequences contextually. The facilitator will lead the discussion and make it as interactive as possible by providing what-if types of discussions. The facilitator must be able to help the trainees to realize the criticality of the stressor and that it poses detrimental consequences. The process should be able to develop a sense of self-realization of the importance of the matter in their own organization context. 	
<p>3) Sense making (impact assessment)</p> <ul style="list-style-type: none"> The trainees will point out the solutions or tools that can help to assess the impact and probability of the scenario if any possible by selecting an action card. The facilitator can then moderate the discussion about their choices 	

Game phase and activity	Illustration
<p>4) Anticipation</p> <p>The trainees can select RRM (risk reduction measures) action cards in order to detect, prevent and/or mitigate the consequences of the scenario studied.</p> <p>They can adapt mitigation measures that are already available in the list of action cards or they can suggest new types of measures used in their specific utility's configuration. If new measures are proposed by the trainees, a card should be created within TORC.</p>	
<p>5) Feasibility</p> <p>Players optionally discuss and assess the requirements (human resources, training, financial resources etc.) to implement the measures selected and the needs within the specific utility and/or the boundary conditions described by the facilitator.</p> <p>The aim of this step is to assess the actual feasibility to implement the selected measures, the expected time at which the actions are expected to be effective and define the requirements to facilitate their adoption.</p>	

Game phase and activity	Illustration
<p>6) Decision</p> <p>Based on the outcomes of the discussion in the previous steps, the most effective solution(s) are chosen by the relevant players with the decision-making roles. The selection might be complemented with the list of requirements to be followed up in order to make the adoption of the feasible solutions.</p> <p>This can be the end of the game unless the facilitator introduces a new stressor within the system at the current state of the game.</p>	 <p>TORC GAME</p> <p>DETECT PREVENT MITIGATE</p> <p>GENERAL VALUE PREVENT AND MITIGATE THE SCENARIO ADAPTATION ANTICIPATE & ENGAGE MANAGE UNDESIRABLE CONSEQUENCES</p> <p>REQUIREMENTS DETECT PREVENT</p> <p>STOP-IT</p> <p>CHOSEN</p> <p>(Human resources, training, investment + ...) requirements for application in (locally adapted) scenario</p>
<p>Continuation (new round) (optional step)</p> <p>Before moving to the "general value" step, the facilitator can decide to run multiple sessions of the game (steps 2-6) by adding additional stressors to the scenario resulting from the decision step of the previous game session. In this case the trainee will start the game over again from step 2.</p>	 <p>What-if?</p> <p>Stressor X</p> <p>Stressor 1</p> <p>Locally adapted scenario</p> <p>Consequences?</p> <p>Criticality?</p> <p>TORC GAME</p> <p>STOP-IT</p>
<p>7) Monitoring and General value</p> <p>The facilitator summarizes with the help of the trainees the findings of the game, pointing out how the adopted measures should be monitored in the long-term perspective.</p> <p>The conclusions may include the recommendations for utility's decision-makers or for being shared with other water utilities. It contains the adaptive action path that the players have adopted during the game regarding the different choices that they made.</p>	 <p>3 VALUE</p> <p>Stressor X</p> <p>Scenario</p> <p>Facilitator name and organisation</p> <p>Actions agreed</p> <p>Notes</p> <p>General recommendations</p> <p>CHOSEN</p> <p>(Human resources, training, investment + ...) requirements for application in (locally adapted) scenario</p> <p>STOP-IT RMD</p>

In the following, two examples of a playing session are provided to help trainers and trainees to create their own TORC sessions.

STOP-IT Case: denial of service due to signal jamming

In this paragraph, we develop an example for a playing session. The scenario focuses on the Denial of Service due to signal jamming [28]. We explore the crucial role of the facilitator in conducting the session and how she/he walks the player through different steps of the game.

Table 6: Example of playing with STOP-IT TORC

Game phase	Activity
1) Framing and setting the scene	<p>After a preliminary talk with the utility's risk assessment officer, the facilitator selects a specific part of the infrastructure (geographically or based on the problematic sectors identified by the risk assessment officers). Then, at the beginning of the game, he/she distributes randomly the role of decision-maker, risk assessment officer or staff in charge of operational activities (to the players). As this utility uses a certain type of telecommunication technology to send and receive data between the control room, cloud, sensors and actuators, a possible stressor is jamming that can cause the assets not to send or receive data to/from the control room.</p> <p>The facilitator describes to the players the sector where the stressor may happen. For example, reservoirs' levels are critical, and the control room wants to activate the pumping stations. However, due to jamming, the command signal is not received by the pumping stations.</p>
2) Awareness	<p>Players discuss about the possible effects of jamming on the infrastructures and its impacts on the service provided. In our example above, the reservoirs will stay below critical level and empty if no action is taken leading to certain number of customers without water supply for a period of time.</p>
3) Sense making (impact assessment)	<p>The players select the action card "Stress-testing platform – module 1" in order to assess the number of the customers and the period of time without water supply based on the KPI.</p>
4) Anticipation	<p>The players select action cards, distinguishing between solutions to prevent the event from happening (e.g., jamming detector) and/or to mitigate the consequences (e.g., increasing the redundancy of the physical and / or IT system).</p> <p>Building on our example, the players might sort out possible action cards such as the jamming detector, cable connection, manual operation of the pumps, etc.</p>
5) Feasibility	<p>For each of the action cards proposed, the players discuss and assess the resources needed and the requirements to implement each option.</p> <p>In our example, they should evaluate, for instance, the need of additional manpower for the manual operation of the pumps, or the investments for cable connection, or the price for buying and implementing a jamming detector.</p>

Game phase	Activity
6) Decision	The players with decision making role, after considering the information discussed at the feasibility stage, choose the best option of mitigation measures to be implemented.
7) Monitoring and General value	The facilitator describes the decisions made by the players and summarizes the effect of the main solutions adopted on the long run, while providing recommendations to be followed up by the organization.

DWC Case: spoofing of a web application impacting the service

In this paragraph, a second example for a playing session is provided and inspired by a use case of the DWC project developed in the Risk Guide [27]. The Risk Guide provides an example on how to apply the Risk Management Procedure (RMP) according to ISO 31000:2018 in one of the digital water cities, i.e., the city of Copenhagen. Specifically, a cyber-attack of the Digital Solution 13 (DS13), developed by the Danish project partners BIOFOS and DHI, have been considered to show the applicability of the Risk Guide in a real case of the water domain. This case-study has been adapted to certain selected phases of the TORC game in order to facilitate the main concepts contained in the Risk Guide to perform a proper RMP.

Table 7: Example of playing with TORC for a DWC case

Game phase	Activity
1) Framing and setting the scene	The facilitator tells the players about the background of the scene which will be played. The hacked digital solution and the characteristics of the considered Wastewater Treatment Plant (WWTP) are described. Moreover, the description of the "attack", which is known as the "stressor" in the TORC game, is provided without quantifying the estimated impact that such event could have on the system. Finally, relevant Key Performance Indicators (KPIs) and risk criteria adopted by the organisation to limit the pollution of the receiving water body are described and explained to the audience.
2) Awareness	At the beginning of this phase the players can ask questions to the facilitator to have a more detailed picture of the stressor and the system they are playing with, in order to achieve a complete identification of the risk. The discussion should be also about the understanding of what can happen if the considered digital solution is hacked. Afterward, the facilitator sets the ground to a complete description of the risk, from the source to the type of impact. In the specific case the nature of the loss is related to the environment and to the reputation of the involved organisation.
3) Sense-making	First, the decision maker asks support to Risk Analysis experts which are capable of figuring out how to evaluate probabilities and consequences of the identified risk. Then, the facilitator explains the suggested stress-testing procedure used for the case-study for the quantification of consequences, without asking the trainees to run a simulation model because this would be out of the scope of the specific training activity. After directly providing the final results of the stress-testing, in terms of the selected KPI, the facilitator introduces the Infrarisk-CP

Game phase	Activity
	assessment which can be interactively carried out, though the use of an Excel spreadsheet, provided before or during the game session. Thus, the players can get familiar with the concepts of exposure and vulnerability which are intrinsically connected with the likelihood of having a successful cyber-attack.
4) Anticipation	Here, the "Risk Treatment" step of the Risk Guide is performed. In fact, after combining the estimated consequence and probability, the risk is calculated and evaluated on the basis of the risk criteria which were set at the beginning of the game. According to the actual Level of Risk, the players can select different action cards, representing different mitigation measures which can be adopted. If risk is considered unacceptable, the players discuss which measures could be applied, either to reduce probability or consequence. Expert of WWTP and IT are called to propose risk treatment options not yet in place. The mentioned action cards can be inspired by the Risk Mitigation Measures Database (RRMD) which is described in the mentioned step of the Risk Guide called "Risk Treatment".
5) Feasibility	At this stage the players are eventually asked to calculate the benefits on the systems of the selected potential risk mitigation measures. In the provided Infrarisk-CP spreadsheet, different measures have an impact on the probability of success of the considered attack so the different alternatives can be compared, also discussing about the potential costs of each solution.
6) Decision	Here, the facilitator asks the players to take the actual decision, having the decision maker invited to select the most feasible alternative (or set of alternatives) among the selected action cards which are supposed to mitigate the risk. A consultation with all the other played is recommended.
7) Monitoring and General value	A final discussion is triggered by the facilitator about the potential options for a proper monitoring of the situation along the time (after risk reduction measures have been implemented) and about the value creation achieved during the game session.



Leading urban water management to its digital future

digital-water.city
 **digitalwater_eu**



digital-water.city has received funding from the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 820954.