D2.6 Final Ethics and Legal Framework

Work Package:	WP2		
Lead partner:	Sheffield Hallam University (CENTRIC)		
Author(s):	Helen Gibson (CENTRIC), Rowan Dennis (CENTRIC)		
Due date:	30 September 2022		
Version number:	1.0	**Status:**	Final
Dissemination level:	PU: Public		

Project Number:	883284	**Project Acronym:**	7SHIELD
Project Title:	Safety and Security Standards of Space Systems, ground Segments and Satellite data assets, via prevention, detection, response and mitigation of physical and cyber threats		
Start date:	September 1st, 2020		
Duration:	30 months		
Call identifier:	H2020-SU-INFRA-2019		
Topic:	SU-INFRA01-2018-2019-2020 Prevention, detection, response and mitigation of combined physical and cyber threats to critical infrastructure in Europe		
Instrument:	IA		

Horizon 2020
European Union Funding
for Research & Innovation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883284

Revision History

Revision	Date	Who	Description
0.1	02/09/2022	CENTRIC	ToC and Initial Content
0.2	16/09/2022	CENTRIC	Update to General Legal Framework
0.3	23/09/2022	CENTRIC	DPIA and remaining sections
0.4	26/09/2022	CENTRIC	Document ready for internal peer review
0.5	30/09/2022	CENTRIC, CLS, RG	Update with review comments and finalisation
1.0	01/10/2022	CENTRIC, ENG	Final version

Quality Control

Role	Date	Who	Approved/Comment
Internal review	28/09/2022	CLS	Approved with minor edits
Internal review	30/09/2022	RG	Approved with minor edits

Disclaimer

Executive Summary

7SHIELD (Safety and Security Standards of Space Systems, ground Segments and Satellite data assets, via prevention, detection, response, and mitigation of physical and cyber threats) is aimed at addressing the security and resilience of Ground Segments (GS) of Space Systems. Such systems provide enormous amounts of critical satellite data for earth observation (EO), satellite communications (SATCOM) and global navigation satellite systems (GNSS). European citizens, public sector and commercial sector services all depend on access to these services every day.

7SHIELD is an extensive prevention, detection, response, and mitigation system that aims to provide tools that support the operators of satellite ground segments throughout the crisis lifecycle. The system provides functionalities that operate during the pre-crisis, crisis, and post-crisis phases. Such tools and technologies; however, require careful oversight to ensure that they effectively operate in a legally and ethically compliant manner implementing the appropriate safeguards to protect the fundamental rights of all European citizens.

This is the second of two deliverables focused on the development of an ethical and legal framework to underpin 7SHIELD as an operational system. The first, D2.3 Preliminary Ethics and Legal Framework, provided an extensive overview of the underlying legislative environment as well as their applicability to each of the 7SHIELD technologies alongside any corresponding ethical implications. The locations of the five pilot sites were also considered to understand any specific implications of any national legislation.

Therefore, in this deliverable we focus on providing updates based on advances and introductions of various new legislative controls and environments since the submission of D2.3 in February 2021. In the last 18 months we have seen the finalisation of the United Kingdom's exit from the European Union, The introduction of the proposal for a new Artificial Intelligence Act and progress along the legislative train for the CER Directive on the Resilience of Critical Entities and the update to the Network and Information Security Directive, the so-called NIS2. Thus, first, we provide updates to the legal context on these three pieces of legislation and their applicability to 7SHIELD.

Secondly, we reassess the 7SHIELD technologies, divided into prevention, detection, response and mitigation technologies, and provide any further updates to the corresponding legislation as well as how they may be impacted by the AI Act.

Finally, we carry out a preliminary data protection impact assessment of the 7SHIELD system focusing on the available technologies and their use of personal data to understand the full scope and potential impact of the system, and to highlight where further safeguards could need to be put in place to ensure safety and viability in an operational environment.

Table of Contents

List of figures

Definitions and acronyms

AI	Artificial Intelligence
CA	Consortium Agreement
CCTV	Closed Circuit Television
CER	Critical Entity Resilience
CI	Critical Infrastructure
CIP	Critical Infrastructure Protection
CNIL	Commission Nationale de l'Informatique et des Libertés
C/P	Cyber/Physical
CTIP	Cyber Threat Intelligence Platform
DPIA	Data Protection Impact Assessment
EC	European Commission
ECI	European Critical Infrastructure
EECC	European Electronic Communications Code
EO	Earth Observation
EU	European Union
EUCI	European Union Classified Information
FDR	Face Detection and Recognition
GDPR	General Data Protection Regulation
GNSS	Global Navigation Satellite System
GS	Ground Segment
HLEG	High level Expert Group on AI
MS	Member State
NIS	Network and Information Security
ODE/AR	Object detection and activity recognition
SATCOM	Satellite Communication
SGS	Satellite Ground Station
SSO	Single Sign On
TDSS	Tactical Decision Support
UAV	Unmanned Aerial Vehicle
UN	United Nations
WP	Work Package

1. Introduction

7SHIELD (Safety and Security Standards of Space Systems, ground Segments and Satellite data assets, via prevention, detection, response, and mitigation of physical and cyber threats) is aimed at addressing the security and resilience of Satellite Ground Segments (SGS) of Space Systems. Such systems provide enormous amounts of critical satellite data for earth observation (EO), satellite communications (SATCOM) and global navigation satellite systems (GNSS). European citizens, public sector and commercial sector services all depend on access to these services every day.

The safety and security of SGS are essential as a critical infrastructure (CI) themselves but the data they provide are also used for monitoring other CI sites and for supporting emergency response in the event of a major disaster. Such sites must have adequate protection and resilience to prevent and respond to both natural disasters and man-made attacks that impact the physical environment and cyber operations within the ground segment.

To protect such infrastructure, 7SHIELD has developed a comprehensive system that provides a range of technological components and a series of pilot demonstration events that test how such a system would be deployed and used operationally. An important consideration is the legal and ethical framework in which the operational deployment of such advanced technologies would exist.

This deliverable builds on and provides an update to D2.3 Preliminary ethics and legal framework that was submitted at M6 of the project. The two deliverables should be seen as complementary to each other. The first, D2.3, provided the foundational information, scope and context for the legal and ethical assessment that will not be repeated in detail here. Instead, this deliverable takes the following approach. Firstly, several new pieces of legislation have either been introduced or have advanced on the legislative train since the production of D2.3. The most critical of these are:

- the proposed introduction of the Artificial Intelligence Act,
- the Directive on the Resilience of Critical Entities (CER Directive), and
- the advancement of the NIS2 Directive. These are presented in Section 2.

Furthermore, the extent to which each 7SHIELD technology is defined and understood is now significantly more advanced, therefore an enhanced conversation on any potential legal or ethical considerations can also be held (Section 2.4) whilst there is little to update in terms of national legislation until the above EU directives pass through the European Parliament and need to be transposed into national legislation. Therefore, although we expected some updates to the national legislation it appears that there are limited changes at this stage and major updates will come once the upcoming directives/regulations are

adopted and there is a better understanding on how they will be interpreted into national law.

Finally, as D2.3 proposed to conduct a data protection impact assessment (DPIA) of the overall 7SHIELD system, in the final section (Section 4) we go some way to achieving that goal (whilst ensuring that no confidential or European Union Classified Information (EUCI) is divulged within the discussion). The final section then summarises the outcomes and identifies any future considerations required to bridge the gap to an operational system/deployment.

2. Updates to the general legal framework

In D2.3 Preliminary Ethics and Legal Framework, we covered the core legal principles that apply to the 7SHIELD project, this included:

- Overview of Fundamental Rights and their underpinning applicability across the legal frameworks of the EU and UN.
- Application of the General Data Protection Regulation (GDPR) as the prevailing piece of legislation within the EU regarding the protection of personal data.
- Specific legislation relating to the domain of 7SHIELD focusing on
    - Protection of national critical infrastructure,
    - Space and space ground segments,
    - Cybersecurity.

In this section, we provide update on the above areas and how they may have potential implications on the future operational deployment of the 7SHIELD system. Specifically, we cover the following updates:

- *(GDPR) Finalisation of the exit of the United Kingdom of Great Britain and Northern Ireland from the European Union – incorporation of the GDPR into UK law and the adequacy decision between the European Union and the UK.*[1]
- *Additional Legislative controller added relating to Artificial Intelligence – called upon by the European Union (Proposal 2021/0106) – Artificial Intelligence Act*
- *Updated Legislative controller added relating to the Resilience of Critical Entities – called upon by the European Union (Proposal 2020/0365)*
- *Updated Legislative controller added relating to the NIS2 Directive – called upon by the European Union* (COM/2020/823).

2.1. Data and the General Data Protection Regulation

The GDPR as a regulation has remained the same since the D2.3; however, given the finalisation of the United Kingdom's (UK) withdrawal from the European Union updated legislation now applies between the UK and the EU. Overall, this is not a significant concern within the scope of this deliverable and the testing and operation of the 7SHIELD system as this does not take place in the UK. Nonetheless, it is important to note that the GDPR has been transposed into UK law as part of the Data Protection Act 2018 and is now referred to as 'UK GDPR.' The EU and the UK have reached an adequacy decision on the protection of personal data by the UK in June 2021. This affirms that current UK law provides for

[1] Trade and Cooperation Agreement between the European Union and the European Atomic Energy Community, of the one part, and the United Kingdom of Great Britain and Northern Ireland, of the other part. OJ L 149, 30.4.2021, p. 10–2539. Available at: http://data.europa.eu/eli/agree_internation/2021/689(1)/oj

adequate protection of personal data from EU within the UK.[2] Currently there are numerous discussions within the UK on potential updates to the UK GDPR. In the event of legislative changes, it would be necessary for the EU to reaffirm that the adequate protection remains allowing the UK to retain its status.

In the area of data protection, the proposal for a new Data Governance Act to support data sharing could also become relevant to 7SHIELD in the future. The proposed Data Act[3] provides for more standardisation within datasets supporting the sharing of information between sectors, especially in the manner that can support the future development of Artificial Intelligence applications.

2.2. Artificial Intelligence Act (forthcoming)

The recent developments in the European Union in the year of 2021/2022 has seen a rapid growth in the perception of technology and how it can and should be applied lawfully within the European domain. In April 2022, Ursula Von der Leyen (President of the European Commission) stated in a press release entitled "The Declaration of the Future of the Internet" the vision of:

"…the Internet a safe place and trusted space for everyone, and to ensure that the Internet serves our individual freedom. Because the future of the Internet is also the future of democracy, of humankind."

This statement brought back into the public focus the dynamic environment that is digital technology and the need to ensure such technology is harnessed and regulated in a manner that benefits all. Another example of such technology is the growing presence of Artificial Intelligence (AI). This too is a powerful technology that has extensive possibilities for improving and accelerating many activities that rely and use large amounts of data to make decisions, but it is recognised that it needs to be applied in such a manner that promotes, protects, and safeguards the fundamental freedoms and rights of European citizens.

Therefore, the implementation of a harmonised set of rules to support the adoption, uptake, and implementation of technology, can continue to demonstrate an active stance of protecting the internet and ensuring that technological progress does not diminish a safe and trusted digital space within Europe. Therefore, the same rights and freedoms that are preserved under legislation such as the GDPR should also be enshrined through other legislation – for example, AI. As seen with the GDPR, Human Rights must ensure that the rights and privacy of data subjects whose data is utilised withing AI systems, are also not further subjected to any infringement of their basic rights.

[2] Commission Implementing Decision of 28.6.2021 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate protection of personal data by the United Kingdom CENTRIC (https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/brexit_en)

[3] European Commission (2021) European Data Governance Act. https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

These factors have also brought around a proposed **Coordinated Plan for Artificial Intelligence**[4] which identifies the key areas that provide a level of harmonisation around the use of Artificial Intelligence within the social-economic climate of Europe. This section will dissect the new recommendations (including the coordinated plan) in how the European Union aims to achieve their goal of creating a more harmonised perception for AI in all areas of operation. We will then specifically focus on the 7SHIELD domain and on areas of technologies within the remit of 7SHIELD and the space sector.

In the previous deliverable the overview of AI focused on the Ethics Guidelines for Trustworthy AI,[5] which has now evolved into the above-mentioned Coordinated Plan and will ultimately culminate in the realisation of the proposed Artificial Intelligence Act. The following sections will dissect the proposal for the AI Act and the corresponding requirements for Artificial Intelligence implementation including how it may impact upon the 7SHIELD system now and in the future.

2.2.1. Artificial Intelligence Act: impact, approach and objectives

The new Artificial Intelligence Act launches a new approach to the implementation process of AI systems in any industry or organisation; however, in this document we tailor our interpretation specifically towards the legislative environment of 7SHIELD. The new act proposes a consolidated method for developers, integrators, managers, and other personnel to adhere to when implementing an AI system.

One of the key elements of the act, is to lay down a harmonised definition of an artificial intelligence system. Specifically, Article 3 states that an:

> *'Artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.'*

Where we note that the techniques and approaches listed in Annex 1 include the following:

> *(a)* ***Machine learning approaches****, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;*
>
> *(b)* ***Logic- and knowledge-based approaches****, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;*
>
> *(c)* ***Statistical approaches****, Bayesian estimation, search, and optimization methods.*

[4] Europe Commission. (2021). Coordinated Plan on Artificial Intelligence 2021 Review. Technical report, European Commission, Brussels. https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review

[5] EU High-Level Expert Group on AI (2019) Ethics Guidelines for Trustworthy Artificial Intelligence. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

This highlights that AI, in the context of the proposed AI Act includes a broad range of approaches ultimately meaning the act will cover an extensive scope of solutions, many of which may already in operation and may have been for many years.

The main goals for the new AI Act are to ensure that the standing legislation and fundamental freedoms are not threatened by new technological enhancements and that these rights are safeguarded appropriately and in a harmonised manner. To achieve this goal, the AI Act provides four objectives for systems to meet which will ensure that the overall goal of achieving a safe, robust, and lawful Artificial Intelligence system.

The objectives which the AI Act aims to successfully amalgamate are (i) safe and respecting existing law on fundamental rights and union values. Artificial Intelligence systems must also (ii) ensure legal certainty, (iii) enhance governance and effective enforcement of existing law on fundamental rights and safety and finally (iv) facilitate the development of a single market for lawful, safe, and trustworthy AI applications.[6]

The European Union (EU) aims to implement this by implementing a 'risk-based' system which constitutes how much of an impact the usage Artificial Intelligence would have on a specific critical entity, data subject or sector within a European Union Member State (MS). The risk-based approach is presented in a tiered system which increases the level of risk based on the possibility of infringing upon a user's rights or the inability to coincide with the current enacted legislation of the MS. The four types of risks that are detailed within the Draft AI Act are stated as:

- *Unacceptable Risks*[7]
- *High Risks*[8]
- *Low Risks*[9]
- *Minimal Risks*[10]

Unacceptable risks have been identified within the Draft AI Act under Title II as a Prohibited Artificial Intelligence Practice.[11] Unacceptable Risks are defined as the systems that can be found to 'contravene union values for instance by violating Fundamental Rights.'[12] The European Union through Article 5 has unequivocally defined the parameters for what can be constituted as Artificial Intelligence that goes against the European Union. The developments in what is deemed acceptable will incorporate the Trustworthy AI which guidelines have been produced by HLEG in 2019. The new act has brought 'two years of

[6] Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM (2021) 206 final, 21 April 2021
[7] Art 5 Draft AI Act
[8] Art 5 Draft AI Act
[9] Art 5. Draft AI Act
[10] Art 5. Draft AI Act
[11] Art 5. Draft AI Act
[12] Art 5. Draft AI Act

analysis and close involvement of stakeholders'[13] including 'preparatory work'[14] to bring AI 'that is guided by certain essential values value-oriented principles.'[15] These principles include improving documentation[16] and traceability[17] plus providing more information to the user[18] that holds credibility. Human Oversight[19] techniques should also be utilised to ensure that the highest standard of robustness[20], safety[21] and cybersecurity[22] are operating in an accurate[23] manner to the stated objectives.

2.2.2. Artificial Intelligence Act 2021 and 7SHIELD

The 7SHIELD project utilises a range of Artificial Intelligence systems which would be required to coincide with the new AI Act when it comes into force. While the legislative train is not complete and several areas of the Act are still under discussion, it is prudent for an AI system currently in development or use to be aware of and align as much as possible with the provisions within the proposed act.

The provisions that are present would not prohibit the functionalities of Artificial Intelligence being employed within 7SHIELD; however, the some of the current systems may require further scrutiny and justification of their use and implementation according to the guidelines. The question turns to the factors relating to how AI should be used in the remit of space and critical entities. The primary focus of 7SHIELD is to improve the security of these areas without the loss of protection to core fundamental rights and data privacy requirements. The incorporation of AI in 7SHIELD supports the improved efficiency of dealing with cyber and physical security related incidents through a range of technological solutions that are discussed in more detail in Section 2.4. The impact of the AI Act will ensure the developers of these instruments and their implementation into practice in Europe via 7SHIELD are aware of the new audit requirements in determining the risk of such a technology.

The movement away from uncertainty in this field and into a core citizen and risk-based approach to develop the understanding of how AI works is designed to allow more powerful techniques to be utilised to support cybersecurity and ensure the protection of critical infrastructure.

[13] Art 3.2 Draft AI Act
[14] Art 3.2 Draft AI Act
[15] Art 3.2 Draft AI Act
[16] Article 11 Draft AI Act
[17] Article 12 Draft AI Act
[18] Article 13 Draft AI Act
[19] Article 14 Draft AI Act
[20] Article 15 Draft AI Act
[21] Article 15 Draft AI Act
[22] Article 15 Draft AI Act
[23] Article 15 Draft AI Act

The scope of the act has been developed to align how AI technology is implemented into a functioning society whilst ensure who are exposed to the outcomes of AI systems are protected according to the fundamental rights and freedoms of European citizens.

Depending on the methods used and the complexity of the desired objective for the AI system, as stated above, alters the extent to which raise the risks associated with using them. Such risks can include the likelihood of errors or degree of uncertainty, biases, privacy and more. Similarly, the impact of so-called 'poor' outcomes from AI can differ depending on the type of outcome desired. Therefore, the risk-based approach ensure that the appropriate risk assessment and scrutiny has taken place to be able to ethically and ultimately lawfully utilise AI within the European Union.

7SHIELD deploys a range of systems, some of which utilise these technologies across the project. Given the relatively broad definition of AI within the Act many technologies will ultimately fall under its scope. Due to the range of areas (Space, Ground Segments, Critical Infrastructure (CI), and cybersecurity) industries upon which 7SHIELD is a stakeholder, the protection of society, rights and freedoms including GDPR is a core requirement, therefore, the risk exposure should be assessed accordingly using the risk-based approach. The following sections explain the requirements for determining the level of risk within an AI system and what that means for implementation.

2.2.3. Risk Identifiers

Risk Identifiers have been introduced within the Draft AI Act to provide a functional auditing system that will ultimately be enshrined within EU legislation. The primary source of audit through this act is the risk identifiers highlighted below which dictate what is deemed as an acceptable or unacceptable level of risk and the measures that should be applied to mitigate against such risks.

2.2.3.1. Unacceptable Risks

AI systems that are prohibited are deemed to have unacceptable risks. These have been identified as systems that operate using subliminal techniques or exploit a vulnerability of a specific group of people and therefore, ultimately could cause physical or psychological harm. In addition, AI that performs social scoring that could result in detrimental treatment is also prohibited by the act alongside the use of real-time remote biometric identification in public spaces for the purpose of law enforcement unless it falls specifically under one of the following exceptions (as listed in Article 5(d) of the draft act):

(i) *the targeted search for specific potential victims of crime, including missing children*

(ii) *the prevention of a specific, substantia, and imminent threat to the life or physical safety of natural persons or of a terrorist attack*

(iii) the detection, localisation, identification or prosecution of a perpetrator or suspect of a crime with a maximum sentence of at least 3 years that would allow for the issuing of a European Arrest Warrant

Therefore, we expect that 7SHIELD should not directly implement such technologies unless they fall under one of the exceptions.

2.2.3.2. High-Risk Systems

High-risk systems are not forbidden or prohibited within the EU under the Act but are permissible only if they are developed within the best interests of Chapter 2, Title III of the act which provides mandatory requirements for the user to mitigate the risks that could result in threats to health, safety and fundamental rights.[24] There are seven requirements of the service providers established within the act, and are as follows:

(1) the operation of a *risk management system*,[25]

(2) the use of *high-quality datasets*,[26]

(3) the establishment of *appropriate documentation*,[27]

(4) the inclusion of *logging capabilities* to enhance traceability,[28]

(5) the sharing of *adequate information* with the end-user,[29]

(6) the design and implementation of *appropriate human oversight measures*,[30] and

(7) the achievement of the highest standards in terms of *robustness*, *safety*, *cybersecurity*, and *accuracy*.[31]

Examples of systems that can be considered as high-risk include non-law enforcement uses of real and post remote biometric systems, safety components in the operation of CIs, for access to/assessment for education and training, recruitment and selection in relation to employment, for determining access to public benefits/services including credit scoring, LEA use for AI for assessing (re-)offending risk, polygraphs, deep-fake detection, evidential reliability, criminal profiling, crime analytics and pattern matching, in migration and border control for polygraph/emotional assessment, risk, document validation, residence and visa applications, and judicial assessment of facts.

Due to the nature of a high-risk AI system, the Draft AI Act now also proposes these specific types of systems will need to comply with a mandatory EU Declaration of Conformity[32] which includes the following obligations:

[24] Recital 43 and Art. 7(2) Draft AI Act.
[25] Art. 9 Draft AI Act.
[26] Art. 10 Draft AI Act.
[27] Art. 11 Draft AI Act.
[28] Art. 12 Draft AI Act.
[29] Art. 13 Draft AI Act.
[30] Art. 14 Draft AI Act.
[31] Art. 15 Draft AI Act.
[32] Art 48 Draft AI Act.

- implementation of a *quality management system* in accordance with Art. 17 Draft AI Act,[33]
- provide the *technical documentation* of the *high-risk AI system*,[34] and
- keep the *logs* automatically generated by the *high-risk AI system*.[35]

Furthermore, it has been proposed that such systems will need to be registered in an EU Database for high-risk systems.[36] This is determined under Article 60 of the Draft AI Act, it is a step that must be carried out before the tool or system can be placed on the market (for sale) or utilised.[37]

The Draft AI act calls for more human oversight in AI systems, therefore, it proposes that the obligations of the parties who are utilising the system are only to do so in the accordance of the instructions of the system provider pertaining to Article 29(1). This is designed to ensure that the users of the software are aware of the risks; Human Oversight is not an obligation, however, left to user discretion.

2.2.3.3. Low and Minimal Risk Systems

Protecting the fundamental freedoms and the basic laws of EU MS; especially the rights to privacy, is at the forefront of the Draft AI Act, therefore, low and minimal risk systems also contain obligatory requirements to satisfy which ensure that the protection of data subjects continues across Europe. The first transparency requirement for these systems refers to the providers of the AI system who must inform their data subjects that their system intends to interact with them (a natural persons). The following requirements from Title IV of the Draft AI Act state that information must be relayed to a natural person if the AI system performs any of the following tasks:

- interact with humans,
- are used to detect emotions or determine association with (social) categories based on biometric data, or
- generate or manipulate content ('deep fakes').

Exemptions however can be found in specific settings such as Law enforcement and in cases of Freedom of Expression. The Artificial Intelligence act calls for these requirements based around the abilities for the data subject to omit and make an 'informed choice or step back from a given situation.'[38]

[33] Art. 16(b) Draft AI Act.
[34] Art. 16(c) Draft AI Act.
[35] Art. 16(d) Draft AI Act.
[36] Article 60 Draft AI Act
[37] Article 51 Draft AI Act.
[38] Title IV Draft AI Act.

2.2.4. Future of the AI regulation

The AI Act is currently within its proposal stages and was originally proposed to the European Commission in April 2021. The regulation is currently undergoing MS approval, therefore, recommendations to alter the draft have been presented by the Slovenian Presidency who have requested that specific definitions are to be changed on how AI systems are presented. Further recommendations have been presented by the committee on Civil Liberties, Justice and Home Affairs and Committee on Internal Market and Consumer Protection who have created a collective report. The scrutinization of regulation provides a plethora of opinions within a democratic European Union.

The date of enactment has not been confirmed yet, however, once voting on amendments is expected in the second half of 2022 with the hope that the act will pass into law during the Swedish presidency in the first half of 2023. In the next section we highlight some of the elements that have been raised for discussion and potential alteration before the act is finalised.

Compromise AI Draft of the Council's Slovenian Presidency

The Slovenian Presidency displayed several concerns around the definition of AI and the usage of high-risk AI. The product of these concerns was a '*Compromise AI Act*' which suggested that the definition of AI Systems that should be altered to the following:

(i) receives machine and/or human-based data and inputs,

(ii) infers how to achieve a given set of human-defined objectives using learning, reasoning, or modelling implemented with the techniques and approaches listed in Annex I, and

(iii) generates outputs in the form of content (generative AI systems), predictions, recommendations, or decisions, which influence the environments it interacts with.[39]

The recommendations have been made to provide a more coherent and explicit definition of AI systems. The recommendations from the Slovenian Presidency also removed the defining statement that AI Systems used by LEAs for *crime analytics* regarding a natural person. This does not exempt LEAs, however, from the obligations associated with high-risk systems. There are systems that an LEA can utilise that the Slovenian Presidency still define as a risk to utilise. In the instances of 7SHIELD and the benefits that can be applied through these recommendations of changes can be found in Article 2(3) which also highlights that AI systems developed exclusively for national security purposes are not to have the regulation applied. Article 2(6) and 2(7) which give leeway to scientific research and development:

[39] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text. Available at: https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf

- This Regulation *shall not* apply to AI systems, including their output, specifically developed, and put into service for the sole purpose of scientific research and development.
- This Regulation shall not affect any research and development activity regarding AI systems in so far as such activity does not lead to or entail placing an AI system on the market or putting it into service.

The suggestions that have been provided by the Slovenian Presidency, therefore, can be found to benefit the 7SHIELD project. The approvement, however, of those changes will come to fruition upon full enactment of the act.

European Parliament Draft Report

The Report produced as part of the ordinary legislative procedure contains 309 Amendments to the proposed AI Act from the European Parliament[40]. These recommendations change the language and areas that require further developments or additions to ensure that the enactment of the Artificial Intelligence Act into European society is a process that is not convoluted. The approvement of these recommendations to be imported into the main Artificial Intelligence Act will therefore then expand the legislative process towards its final phases.

2.3. Resilience of Critical Entities (CER) Directive

The new Directive on the Resilience of Critical Entities is also now in the final stages drafting whereby it is expected to be adopted by the European Parliament in November 2022. The developments of the Critical Entity Resilience (CER) Directive have been occurring regularly over the past 18 months with latest position of the Council of Europe as per the 28th of June 2022 stating that the work required to provide the final text on Critical Entities is currently in the 'technical level to finalise the provisional agreement.'[41] The information provided within the CER Directive provides a new revised perspective on a MSs' interaction with their critical entities. The core functions of the CER Directive are to expand the ten sectors that are determined as the founding factors of critical entities within the European Union. The sectors have been identified as:

- Energy
- Transport
- Banking
- Financial Market Infrastructures

[40] Draft Report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM2021/0206 - C9-0146/2021 - 2021/0106(COD)). Available at: https://www.europarl.europa.eu/doceo/document/CJ40-PR-731563_EN.html

[41] Council of Europe (2022) Press Release. EU resilience: Council presidency and European Parliament reach political agreement to strengthen the resilience of critical entities 28th June 2022<https://www.consilium.europa.eu/en/press/press-releases/2022/06/28/eu-resilience-council-presidency-and-european-parliament-reach-political-agreement-to-strengthen-the-resilience-of-critical-entities/>

- Health
- Drinking Water
- Wastewater
- Digital Infrastructure
- Public Administration
- Space[42]

The identified areas from the CER Directive which determine what can be defined as a critical entity affects 7SHIELD due to the space domain being in scope. The articles of *Proposal 2020/0365* determine the requirements of MS who operate critical infrastructures (CIs) and therefore how the uptake of the work within 7SHIELD could be impacted within the MSs that have ratified and abide by the requirements of CER articles. Therefore, in this section we continue to review and analyse the updates to the CER directive over the past 18 months highlighting the implications for the 7SHIELD project based on any of the updated recommendations provided by the European Parliament and the Council of Europe.

The initial CER proposal was covered within D2.3, however, with the timeline to enactment closing the analysis of the scope towards 7SHIELD and the possible implications of the new requirements of member states across the European Union an update is necessary in this deliverable. In the cyberspace it should be noted that the CER Directive does not supersede the Network and Information Security (NIS 2) Proposal, as discussed in Section 2.4, and highlighted in Article 1(2):

> *This Directive shall not apply to matters covered by Directive (EU) XX/YY [proposed Directive on measures for a high common level of cybersecurity across the Union; ('NIS 2 Directive')], without prejudice to Article 7.*

Any instances that refer to any of the issues regarding cybersecurity shall be a matter under the Network and Information Security 2 (NIS2) Directive. The response that is tailored towards a cybersecurity threat to a critical entity should stem from the enacted requirements of the NIS2 due to the close alignment and synergies with the proposed NIS2 Directive.[43] The following sections will detail how the scope of the CER effects the 7SHIELD to ensure that compliance can be found across the entirety of the project and any future system implementation.

2.3.1. Scope of the Resilience of Critical Entities (CER) Directive

The defined scope of the Directive is found within Article 1(a) stating that the directive has been designed to:

[42] Proposal for a Directive of the European Parliament and of the Council on the Resilience of Critical Entities COM/2020/829 final https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:829:FIN
[43] Proposal for a Directive of the European Parliament and of the Council on the Resilience of Critical Entities COM/2020/829 final https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:829:FIN

"Lay down obligations for Member States to take certain measures aimed at ensuring the provision in the internal market of services essential for the maintenance of vital societal functions or economic activities, in particular to identify critical entities and entities to be treated as equivalent in certain respects, and to enable them to meet their obligations".[44]

The defining factors of this directive can be found to establish new functions for Member States when referring to critical entities. The Directive aims to highlight the gap in the market for a risk-based approach regarding the vital areas of a country's infrastructure; highlighted in D2.3. The CER, however, holds a dynamic scope and holds a set of core elements of a specific infrastructure which would then highlight this as an 'essential service'[45]. The CER functions to identify a risk to critical entities through usage of a risk assessment; a methodology to determine the nature and extent of a risk by analysing potential threats and hazards.[46] A Critical Entity is composed of an 'essential service' which a MS relies upon as a part of their internal national critical infrastructure which if a threat was present it could significantly disrupt the provision of that service it would have impact on other areas of essential services. The criteria demonstrate these requirements in Article 5(2) of the Directive, it shows that:

a) *the entity provides one or more essential services.*
b) *the provision of that service depends on infrastructure located in the Member State; and*
c) *an incident would have significant disruptive effects on the provision of the service or of other essential services in the sectors referred to in the Annex that depend on the service.*[47]

In the final compromised text,[48] the Annex states explicitly that within scope is the space sector and that in particular this applies to: *'Operators of ground-based infrastructure, owned, managed and operated by Member States or by private parties, that support the provision of space-based services, excluding providers of public electronic communications networks within the meaning of point (8) of Article 2 of Directive (EU) 2018/1972'.*

It is a requirement that the identified areas of a MSs' infrastructure that can be determined as a critical entity must be informed. This is due to the requirements of the MS, once it has been identified that there are critical entities present within their country actions must occur to ensure that core fundamentals of society are protected. Detailed under Article 3(1) determines the time frame for member states:

[44] COM/2020/829 Article 1 1(a)
[45] COM/2020/829 Article 2(5)
[46] COM/2020/829 Article 2(7)
[47] COM/2020/829 Article 5 2(a)-(c)
[48] Proposal for a directive of the European Parliament and of the Council on the resilience of critical entities (First reading) - Confirmation of the final compromise text with a view to agreement https://data.consilium.europa.eu/doc/document/ST-12414-2022-INIT/en/pdf

Each Member State shall adopt by [three years after entry into force of this Directive] a strategy for reinforcing the resilience of critical entities. This strategy shall set out strategic objectives and policy measures with a view to achieving and maintaining a high level of resilience on the part of those critical entities and covering at least the sectors referred to in the Annex.

The strategy which the European Parliament has called for takes a pragmatic response to dealing with threats to critical entities which also subsequently demonstrates the specific scope of the new CER directive. Article 3(2) highlights that the strategy should contain:

(a) *strategic objectives and priorities for the purposes of enhancing the overall resilience of critical entities taking into account cross-border and cross-sectoral interdependencies.*[49]

(b) *a governance framework to achieve the strategic objectives and priorities, including a description of the roles and responsibilities of the different authorities, critical entities and other parties involved in the implementation of the strategy.*[50]

(c) *a description of measures necessary to enhance the overall resilience of critical entities, including a national risk assessment, the identification of critical entities and of entities equivalent to critical entities, and the measures to support critical entities taken in accordance with this Chapter.*[51]

(d) *a policy framework for enhanced coordination between the competent authorities designated pursuant to Article 8 of this Directive and pursuant to [the NIS 2 Directive] for the purposes of information sharing on incidents and cyber threats and the exercise of supervisory tasks.*[52]

Article 3 also states that the strategy in question which has emerged from the CER Directive shall also be updated where necessary and at least every four years.[53]

2.3.2. Risk Assessment

Risk assessments within the CER directive function between the member states government and the Critical Entities within themselves. Article 10 states that '*Member States shall ensure that critical entities assess within six months*' and that the assessments must "*account for all relevant risks referred to in Article 4(1) which could lead to the disruption of the provision of essential services.*" The risks that could be opposed are defined in 4(1) as *relevant natural and man-made risks including*:

- accidents,
- natural disasters,
- public health emergencies,

[49] COM/2020/829 Article 3
[50] COM/2020/829 Article 3
[51] COM/2020/829 Article 3
[52] COM/2020/829 Article 3
[53] COM/2020/829 Article 3

- antagonistic threats (including terrorist offences).[54]

The Risk assessments also require MS communication; if, for example, a critical entity is shared across a border this also includes third party countries due to some European states not holding EU status. An example of the requirements of an EU MS communicating with each other from a 7SHIELD perspective refers to any threats or interference to satellite infrastructure that could be utilised by a different MS or nation. An example of this in practice is between the United Kingdom and France with the Eurotunnel ("Chunnel") which spans underneath the English Channel from Folkstone UK (Non-EU) before resurfacing in the French EU State in Calais. It is a requirement of the Critical Entity owner (Getlink)[55] to ensure communications between the French and English controllers of the Chunnel due to the critical nature of the tunnel if an accident was to occur. The CER will call for more inter-member state communications regarding threats to infrastructures. Stakeholders within this transport critical entity between the United Kingdom and France are from other third-party countries, therefore, it must be made clear that a risk assessment would have to be created. The CER directive, therefore, could have a substantive effect on the 7SHIELD project due to the critical entity presence within 7SHIELD. The appropriate risk assessments should be produced by the MS and owners of the specific company that has satisfied the requirements of Article 5(1). which subsequently requires the implementation of Article 3(2) for the entities that are within the EU.

Ultimately, we note that the CER directive confers more obligations on MSs and operators of critical infrastructure than on the providers of systems and software to support the resilience of these entities. Nonetheless, it is important that a good awareness of the CER Directive is maintained across the project and developers to ensure and support compliance in the long term.

2.4. Network and Information Security Directive - NIS2

While the original NIS (Network and Information Security) Directive and the plans for NIS2 were discussed extensively in D2.3, the Council, and the European Parliament as of the 13th May 2022[56] have reached a finalised set of measures for the high common level of cybersecurity across the Union. This is also known as the NIS2 directive.

The text confers additional obligations for the implementation of cyber security in areas covered by the directive. This includes the implementation of appropriate risk management and update reporting obligations for relevant organisations. The directive helps to support

[54] COM/2020/829 Article 4

[55] Getlink are a European Public Company based in Paris which allow the running of DB Schenker Freight trains and the Eurostar company.

[56] European Commission (2022) Strengthening EU-wide cybersecurity and resilience – provisional agreement by the Council and the European Parliament https://www.consilium.europa.eu/en/press/press-releases/2022/05/13/renforcer-la-cybersecurite-et-la-resilience-a-l-echelle-de-l-ue-accord-provisoire-du-conseil-et-du-parlement-europeen/

the harmonisation of cybersecurity requirements, laying the basis of a minimum regulatory framework and, where necessary, the cooperation between MSs.

The following elements of the Directive are particularly relevant to 7SHIELD. As noted above in the review of the CER Directive, the need to harmonisation between the CER and NIS2 Directives is essential, and this is also made clear within the NIS2 Directive alongside the need to ensure this is transposed into national legislation through the means of a cybersecurity strategy. Specifically, Article 5(1)(f) states that the national cybersecurity strategy must include *'a policy framework for enhanced coordination between the competent authorities under this Directive and [the CER Directive] for the purposes of information sharing on incidents and cyber threats and the exercise of supervisory tasks'*. Furthermore, at the EU level the continued cooperation between the groups supporting both the CER Directive and the NIS2 Directive should be realised through regular meeting of at least once per year.

3. Updates to legal and ethical considerations of 7SHIELD technology

In this section, we provide an update on additional legislation or ethical considerations not previously discussed in D2.3 to ensure we have complete coverage of the overall landscape in the scope of 7SHIELD. We follow the same structure as D2.3 by covering in turn the four main elements of the system and the associated technologies developed to address them.

3.1. Prevention Technologies

The prevention technologies are deployed mainly in the pre-crisis phase to assess the extent to which the organisation is prepared for or aware of potential threats that they may be subjected to.

Risk and vulnerability assessments

As discussed in D2.3, the risk and vulnerability assessments are closely aligned with the goals of the CER Directive. In particular, it mirrors its aims in terms of ensuring that operators of critical entities have a good situational awareness of the potential threats they face and, perhaps even more importantly, the extent to which they have an impact on the operation of the CI, the users of its services and depending on the incident also those in the local area.

Any risk and vulnerability assessments that are algorithmically based, also will need in the future to need to bear in mind the scope and implications of the AI act. Such a system would be more likely to fall into the scope of minimal and low risk systems as they do not impact fundamentals rights nor are safety critical systems themselves. Nonetheless, it may be necessary to still ensure that the system is transparent in its operation, especially to users and assessors. This is also an ethical consideration and too much reliance on a single system can also prevent assessors from considering novel threats outside the scope of the system, as well as mechanisms to ensure that the risk modelling and vulnerability assessment stays in line with a changing landscape of threats.

Secure authentication

As mentioned in D2.3, the main consideration for secure authentication is that it supports compliance with the NIS2 Directive. Article 18 of NIS2 describes the cybersecurity risk management measures that should be implemented noting that *'essential and important entities shall take appropriate and proportionate technical and organisational measures to manage the risks posed to the security of network and information systems which those entities use in the provision of their services'*. Therefore, the application of secure authentication mechanisms helps to ensure that CIs can achieve this requirement.

Cascading effects (and the prevention of)

Previously, the need for the consideration of cascading effects was highlighted in the CER Directive, this text remains in the preamble of the final legislation ensuring that not only are the immediate effects of the incident considered but that the ripple effects are taken care of and where possible accounted for in advance of the incident.

3.2. Detection Technologies

It is in the use and application of detection technologies that has probably seen the largest shift in both legal and ethical perceptions since D2.3, in particular, if the proposed AI Act is realised then many of the technologies within this section may fall into the high-risk category and therefore have additional obligations to comply with – especially if such a technology is being developed with a view to bringing the technology to market.

Online data acquisition

Firstly, we consider online data acquisition, specially focused on information accessed and analysed to support threat intelligence. Elements such as data privacy and protection, copyright, terms of service and the robots.txt exclusion protocol. While the general principles of online data acquisition have not changed since the first deliverable, it is important to reiterate that considerations such as terms of service are not a static target but must be regularly reviewed for updates and additional restrictions.

Furthermore, from an ethical and data privacy standpoint, the perception of online content acquisition is constantly moving and updating focus in line with societies' expectation of the extent to which it is acceptable to potentially process reasonable large amounts of data from online sources and whether those who have shared the information on those sites have the same privacy expectations. Furthermore, 7SHIELD is not a law enforcement focused project and there can also be ethical differentials in terms of what a law enforcement body may access versus other public or private bodies may access; even when the goals may be similar. An example of this can be found in the '*HiQ Labs, Inc v LinkedIn Corp. (USA)*' cases where there has been several back and forth on whether it was lawful or not to scrape data from the LinkedIn site.[57]

Video surveillance, processing image and video data

Video surveillance, the use of biometric data for facial recognition and especially its deployment in live or real-time situations is one of the most intensely debated and considered topics over the past 18 months in the legal and ethical space. Previous advice by the European Data Protection Supervisor highlighted both the risks and requirements for utilising such technology but now the potential introduction of the AI act has brought the application of such technologies into much sharper focus.

[57] hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099. Available at: https://plus.lexis.com/api/permalink/9af27d35-b81e-4fd7-8e56-dcf913029077/?context=1001073

7SHIELD technology is not operated by a law enforcement agency and therefore the implementation of such an approach as live facial recognition is not considered at the highest level (i.e., an unacceptable risk) but more likely a high-risk application. This will confer, in the future, several additional rights and responsibilities onto the developers of such technologies to ensure that it complies with the necessary elements of the Act including a specific analysis of the potential risks posed by the implementation of the AI system and the data and governance procedures surrounding the development of the technology.

3.3. Response Technologies

Semantic modelling

We do not foresee any specific updates to the semantic modelling from a legal perspective, nonetheless it is always prudent to keep such technology under evaluation and ensure that the use cases in which it is deployed do not elevate any of the activities into a higher-risk AI system or facilitate the re-identification of persons from disparate personal data.

Wearables and health data

Due to the highly personal nature of the data collected when making use of wearables and other health technology data, it is essential to remain ahead of any privacy concerns. Again, the majority of these either relate to the GDPR and how such data is processed – which should be in a manner that does not negatively affect the personal whose data is collected (either professionally or personally) especially in light of the fact that such information could expose private medical information.

In this case, again it is important to continue to monitor the interpretation of the GDPR and how the data is ultimately processed and used to monitor for possible risks through the AI act and the need to mitigate against these in the future.

Social media communications

The need to inform the public in the event of a critical incident (which can mean an interruption to a service or a threat of harm to them in some form) also forms part of the 7SHIELD system. The need to provide a form of public warning or alerting is catered for in the European Electronic Communications Code (EECC).[58] Article 110 states that where public warning systems are in place the messages can be transmitted via means also other than SMS. While the warning message generation would not necessarily fall directly under the code, such messages may be able to support and amplify official communications. Furthermore, the implementation of the new EU Digital Services Act[59] should also limit the

[58] Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code. Available at: http://data.europa.eu/eli/dir/2018/1972/oj

[59] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC

amount of misinformation online including during crisis events, thus elevating the voice of public authorities and agencies sharing credible and authoritative information.

Ethically, there is also a responsibility on organisations who post of share warning or alerting messages, even with good intention. Sharing of information such as instructions can be particularly beneficial to those navigating a disaster; however, when large numbers of people are involved, the instructions shared may have a direct and consequential impact on their decisions and ultimately their safety. Therefore, ensuring that messages are approved to be shared through official channels and that there is agreed coordination between multiple organisations (where more than one agency may be participating for the response) should ensure that citizens do not receive conflicting advice.

Drone operations

The EU legislation considering drone or UAV operations were covered extensively in D2.3; nonetheless, given this is a fast-moving area of technology EASA (European Aviation Safety Agency) has issued updated guidance on the civil use of drones within the EU. In particular, ED Decision 2022/002/R60 provides updated information for the operation of drones in both the open and specific categories especially regarding the acceptable means of compliance (AMC) elements.

3.4. Mitigation Technologies

Business continuity plans

Business continuity plays an important role in both the CER Directive and NIS2 Directive. In the CER Directive, Article 11 denotes the requirements for operators of critical entities to put in place a fully comprehensive resilience plan that ensures specifically (through 11(1)(d)) to have in place business continuity measures that support their ability to recover from incidents. Many of the elements in the prevention technologies of 7SHIELD also support the obligations under this article. Similarly, Article 18 of the NIS2 Directive also necessitates the development of business continuity measures to effectively manage risk and recover from crisis incidents.

[60] EASA (2022) Executive Director Decision 2022/002/R Amendment 2 to Issue 1 of the Acceptable Means of Compliance and Guidance Material to Commission Implementing Regulation (EU) 2019/947 and to its Annex

4. Data Protection Impact Assessment

Under the GDPR a data protection impact assessment (DPIA) should be carried out whenever there is a high likelihood of data processing operations carrying a high risk to the persons whose data is being processed. Article 35 of the GDPR sets out the terms and requirements for a DPIA, specifically Article 35(1) states:

"Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks."

The assessment itself should cover the following key elements (as set out in Article 35(7)):

- *a systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;*
- *an assessment of the necessity and proportionality of the processing operations in relation to the purposes;*
- *an assessment of the risks to the rights and freedoms of data subjects referred to in paragraph 1; and*
- *the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned.*

Finally, the GDPR itself, along with WP29[61] (the data protection working party) have specifically produced guidance on *'determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679'*. The GDPR provides three broad criteria which are then elaborated by the working party's paper. The core three criteria are:

- *a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;*

[61] WP29 (2017) Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Article 29 Data Protection Working Party. Last revised and adopted on 4 October 2017.

- *processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10; or*
- *a systematic monitoring of a publicly accessible area on a large scale.*

On the strict interpretation of these criteria, the 7SHIELD system as a whole would not perhaps meet the overall definition of a high-risk system; however, when examining the elucidated criteria, it is clear that there is a benefit to carrying out a DPIA under the scope of 7SHIELD. These additional nine criteria are the following:

1. *evaluation or scoring*
2. *automated decision making with legal or similar significant effect*
3. *systematic monitoring*
4. *sensitive data or data of a highly personal nature*
5. *data processed on a large scale*
6. *matching or combining datasets*
7. *data concerning vulnerable data subjects*
8. *innovative use or applying new technological or organisational solutions*
9. *when the processing in itself "prevents data subjects from exercising a right or using a service or a contract"*

It is particularly points 6 and 8 that are relevant to 7SHIELD and thus motivate this DPIA. Previously, 7SHIELD beneficiaries were asked to evaluate whether their activities may result in a need for a DPIA under D9.5 and D9.8. In this case, two partners: SERCO and CENTRIC identified that a DPIA could be required under the scope of their tasks in 7SHIELD. We note two distinctions between this analysis for a DPIA and the analysis in D9.8. Firstly, this assessment concerns the 7SHIELD system, D9.8 is focused on the research activities within the 7SHIELD project. In some cases, this means there are differences in the data processed. For example, CENTRIC identified the need for a DPIA under the social awareness and message generation task under T5.5; however, the DPIA was identified as being potentially necessary for the underlying research into social media rather than for the implementation of the warning message generation component which does not consume personal data – and ultimately the finalisation of the methodological approach focused on organisational communication rather than online communication from individual data subjects. Similarly, some activities within the project occur under a limited and controlled scope or through exemptions that apply specifically to research activities, operation in the wild may necessitate, for components that process personal data, a DPIA. For the SSO, the processing activities were reviewed by the SERCO data protection officer and identified as being not high-risk.

This section is not intended to fully cover every possible mode of the use of personal data within the components but to highlight where a DPIA could potentially be of use and highlight the steps already taken to ensure any impact on data subjects is already safeguarded insofar as possible.

4.1. Methodology

Several approaches to carrying out a DPIA have been published in order to support the process. The article from WP29 suggests the following four frameworks:

- The Standard Data Protection Model published by Germany's Independent National Center for Data Protection[62]
- Guide to Data Protection Impact Assessments by the Spanish Agency for Data Protection[63]
- Guide to Data Protection Impact Assessments by the Information Commissioner's Office (ICO) in the United Kingdom[64]
- Guide to DPIAs by French National Commission on Informatics and Liberty (CNIL) and the associated PIA software and guides[65]

In the assessment we have opted to follow the guidance by CNIL, adapted and structured to support the requirements for 7SHIELD. CNIL sets out a process to support the DPIA that we will describe below.

The CNIL process is divided into four main stages: (1) identifying the context; (2) ensuring compliance with fundamental rights and principles; (3) management of the risks and (4) validation.[66] This approach is mapped onto the following sections. Section 4.2 provides the overall context – i.e., a description of the 7SHIELD system while Section 4.3.1 describes the main processing activities including descriptions of any personal data and associated processes. Section 4.3.3 provides a view on how the fundamental principles apply to the 7SHIELD system, while Section 4.3.3 considers the potential risks and addresses any mitigation measures. We also note this assessment differs slightly from the assessment carried out in D9.8 as here we consider the processing operations for 7SHIELD as an operational system rather than 7SHIELD as a research project.

[62] UAG (2020) The Standard Data Protection ModelConference of the Independent Data Protection Supervisory Authorities of the Federation and the Länder. https://www.datenschutzzentrum.de/sdm/

[63] aepd (2020) Gestión del riesgo y evaluación de impacto en tratamientos de datos personales. Agencia Española de Protección de Datos https://www.aepd.es/es/documento/gestion-riesgo-y-evaluacion-impacto-en-tratamientos-datos-personales.pdf

[64] ICO (2021) Data protection impact assessments. Information Commissioner's Office. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/

65 CNIL (n.d.) Privacy Impact Assessments. Commission Nationale de l'Informatique et des Libertés. https://www.cnil.fr/en/privacy-impact-assessment-pia

[66] CNIL (2018) Privacy Impact Assessment. Commission Nationale de l'Informatique et des Libertés https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-en-methodology.pdf

4.2. Overview of 7SHIELD System

The goal of the 7SHIELD system is to support organizations in the space sector to prevent, detect, respond, and mitigate against a complex variety of physical and cyber threats from a range of natural and man-made scenarios. The system should support all phases of a crisis incident including the pre-crisis, crisis, and post-crisis phases to optimise the prevention and response to any such threats. The system itself is composed of five main groups of technologies:

1. Prevention technologies – focused on risk assessments, threat intelligence and defensive system mechanisms such as secure authentication
2. Detection technologies – focused on data acquisition and processing applications from video, image, thermal cameras, UAV detection and early warning system.
3. Response technologies – focused on further data processing including semantic processing, crisis classification, decision support, message generation and UAV neutralisation
4. Mitigation technologies – focused on post-incident continuity scenarios for both cyber and physical attacks
5. Integration – focused on data models, integration and command and control (C2) system development to access inputs and outputs of the above technologies.

Not all the technologies described above utilise or process personal data thus we only concentrate in this assessment on the elements that will process personal data while any other concerns raised by the technologies are covered in Section 3 on the legal and ethical considerations.

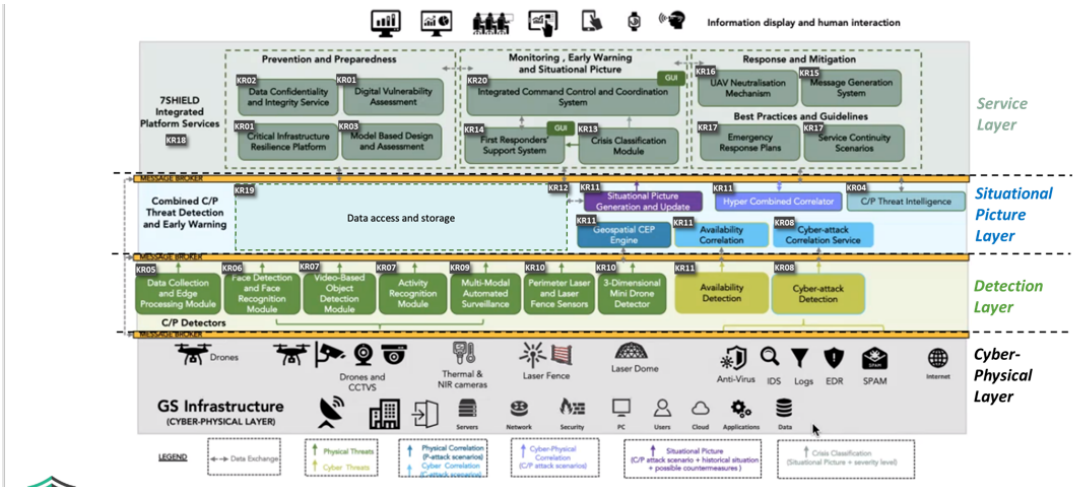

Figure 4-1: Overview of 7SHIELD system

4.3. Data Protection Impact Assessment for 7SHIELD Operational System

4.3.1. Main Processing Activities

4.3.1.1. Prevention Technologies

The goal of the prevention technologies in 7SHIELD are to support the pre-crisis phase. Only a subset of these tools process personal data. The **risk assessment tools** do not process personal data and are there to support CI operators in understanding the risks and consequences of those risks to a part of a holistic risk assessment process. The **single sign-on** (SSO) module allows users of the 7SHIELD system to be created and to login to the various modules using a single set of authorised credentials. The SSO module will process limited personal data (e.g., name, email address, username) to facilitate the access control to the different elements of the system. It is not expected that the SSO element constitutes high-risk processing as it is necessary to maintain the security of the overall system. The SSO prevents brute-force log in attempts as a security feature. Also built into the prevention technologies is a model-based approach to design and assessment of vulnerabilities that models the potential risk features and the cascading effects of multiple risk elements; similarly, this module also will not process personal data.

The cyber-physical **threat intelligence** platform makes up the final component of the prevention tools. The platform aims to support the monitoring of potential physical and cyber threats from multiple sources including from online communities (dark web, social media, marketplaces, forums). The use of such sources indicates the potential for the processing of personal data. The user identifies the sources to be monitored and the natural language processing tools detect whether keywords related to potential cyber or physical threats appear in those sources. The personal data that may appear in the processed data does not form a key part of the data processing action but instead is incidental with the focus on the content of the posts.

4.3.1.2. Detection Technologies

The detection technologies that make up the crisis layer are where the majority of the personal data processing happens within the 7SHIELD system. The threat detection tools include obtaining data from unmanned aerial vehicles (UAVs), intelligence extracted from video surveillance technologies (including face and activity detection and recognition), thermal image processing, laser-based technologies for ground and aerial based intruder detection alongside a cyber-attack detection framework and early warning system.

The data collected from UAVs is processed onboard the UAV using artificial intelligence (AI) techniques and edge processing technologies. Data is collected via camera sensors on the UAV that collect image data during the day, at night and through thermal imaging. The UAV is able to receive instructions that manage the flight path as well as exchanging telemetry data. The cameras on the UAV facilitate the processing of video data by other modules; however, if they detect identifiable persons on the video stream then this could constitute personal data processing.

The **face detection and recognition** module processes personal data in two ways: (1) through the analysis of incoming video streams and still image frames to firstly detect the presence of a face; and (2) by maintaining a repository of authorised faces against which to perform facial recognition based on the video streams received from (1). In terms of high-risk processing the mechanisms it is immediately under consideration as high-risk due to the implementation of live biometric identification technologies. Furthermore, 7SHIELD implements several use cases so the operators of the technology and providers of the video streams can vary across the different use cases.

Similarly, the **object detection and activity recognition** functions also take video streams to identify 'objects' withing multimedia data; this can include inanimate objects such as cars as well as people. If types of activities recognised include actions from people such as walking, standing, running as well as tracking the objects across the video. This detection is also deployed to the UAV so that automated object detection can take place on the edge through the UAV's processing capabilities.

The **cyber-attack detection framework** receives data from various sources including security logs, vulnerability scanning, network traffic and other monitoring to generate alerts for potential security incidents. While the emphasis of this module is not on the processing of personal data, it is possible that the input information could include elements of personal data that have to be managed within the ingestion process.

The use of **thermal and infrared cameras** and imagery can be used to detect persons, animals or vehicles that have a particular heat signature. While the thermal camera can detect a person it is unlikely to be considered as personal data processing in a standalone format given that a specific person is unlikely to be identifiable directly from the thermal image/video stream.

Similarly, the application of **laser technologies to detect persons, vehicles and drones** also has the purpose of detecting the presence of a person but not identifying the person or processing any of their personal data. As with the thermal detection, the location of the person is relayed to the system which could eventually result in their identification if combined either regular cameras or if security on the ground is deployed to that specific location.

Outputs from the described sensors above and processing are then correlated to indicate the presence of an event that the cyber or physical security team will need to react to. The correlation of the data from the sensors is combined to provide indicators such as severity level, assets or areas that are affected and the type of hazard that has been detected. The correlator utilises the results of already processed/aggregated data rather than individual data points; therefore, it does not process personal data itself.

4.3.1.3. Response Technologies

In the context of response technologies, 7SHIELD deploys a **first responder support system** to assist first responders with tactical decision support in the event of an incident that requires a physical presence. The system is able to detect information from physical sensors worn by first responders providing information such as heart rate, temperature, location which is processed and passed back to a C2 unit. Given the data will be received from specific operatives wearing the sensors this will constitute personal and identifiable information, this information will also be viewable by those operating the dashboard. The operatives in the field can also receive information from the C2 dashboard that may include warnings, alerts and instructions as well as supporting communication between team members.

Based on all the incoming data from across the 7SHIELD system, the **crisis classification** module enables the inflight assessment of the severity of risk. This module utilises aggregated and non-personal data to make the assessment. Similarly, the **emergency response plans** allow for users to be guided through the optimal response to an incident based on the specific conditions of the emergency. C2 users access the plans through the dashboard and then work through the guides. The emergency response plans do not exchange data with the rest of the system. Finally, the **message generation system** allows users to construct an automated message that can be disseminated through any available channel to citizens, employees, users, or other groups to inform them about an incident they need to be aware of.

Other response technologies include **UAV neutralisation** that detects and neutralises intruder UAVs that have entered the physical space of the ground segment.

4.3.1.4. Mitigation Technologies

The final set of technologies developed are the **service continuity scenarios** that allow ground station operators to assess the ability of a ground station to continue to operate based on the impact of the incident. The module does not consume or process personal data and therefore does not fall under the DPIA.

4.3.2. Description of Fundamental Principles

Based on the above analysis, although 7SHIELD is an extensive system that brings together large amounts of data, the extent to which personal data is processed within the system (regardless of whether it is high-risk or not) is limited. The main elements that process personal data (regardless of whether the processing is considered high-risk) are the following:

- The single sign-on mechanism (SSO)
- The cyber threat intelligence platform (CTIP)
- The face detection and recognition module (FDR)
- The object and activity detection and recognition module ODE/AR
- The tactical decision support system. (TDSS)

The first stage of a DPIA is to describe the controls guaranteeing the proportionality and necessity of the processing covering the purpose, legal basis, approach to data minimisation, quality of the data collected, and the storage periods.

Purpose

Where personal data is processed it is necessary to demonstrate that it is processed for a specified, explicit, and legitimate purpose and that it is not further processed in a manner that is not compatible with that original purpose. Overall, the personal data is processed within 7SHIELD system to facilitate its correct operation and achieve the stated aim of enabling critical infrastructure to better prevent, detect and respond cyber and physical threats.

SSO – the purpose of the single sign on is to manage the controlled access to the platform. While users have to provide their name and email address this is legitimate for identifying them as a unique user with authority to access to the platform and to allow them to be contacted in case of the need to inform users about changes or outages to the platform (for example) as well as for operations such as validating and resetting passwords.

CTIP – the cyber-threat intelligence platform may collect personal data in the course of obtaining information from online sources; however, this is not the purpose of the data processing and could be considered incidental or collateral. The purpose of the module itself is to ensure that operators are aware of the latest potential cyber threats to important pieces of critical infrastructure.

FDR – the face detection and recognition module will process personal data from streams of video data. The data will be processed from CCTV already deployed in operational sites. The processing is necessary to identify known threats to the ground station site and is not used for speculative processing.

ODE/AR – Object detection and activity recognition is used for highlighting behaviour or unauthorised movement around the ground station site. Although the video data is processed that may lead to the identification of persons, the aim is to analyse behaviour/activity at the aggregated level.

TDSS – The tactical decision support module processes the most personal data within the 7SHIELD system. The goal of the system is to better understand they physical state of the operatives wearing the vests to ensure they remain fit and capable of continuing to operate in the field.

Legal basis

Each processing operation should be underpinned by a legal basis that ensures the lawfulness of the processing operation. The specified legal basis for the overall deployed system can vary depending upon the organisation that the system is deployed at. 7SHIELD has a combination of private and public operators of the space infrastructure; nonetheless, ensuring that the CI is adequately protected is likely to fall under the legal basis of performance of a contract; compliance with a legal obligation; in the public interest or a legitimate interest. In some cases, it is possible that more than one legal basis may apply that guarantees the lawfulness of the processing. Individually, certain modules may also be covered by additional legal basis, for example, the provision of information for the SSO and the operators who opt to wear the vests that provide information to the TDSS should also give their informed consent for the collection and processing of such data

Data minimisation, quality and storage

The data minimisation principle means that only data necessary to carry out the purpose of the processing activity should be collected, no more data than necessary should be stored for a longer period and that all data collected should be accurate. For the SSO, only the name and email address of the user is collected which is necessary to facilitate the services of the SSO. For the CTIP the data collected should be constrained by feeding the system with accurate and targeted sources to minimise collateral collection while in many cases information such as usernames does not need to be retained to competently extract the relevant threat intelligence data. For the FDR, it is important to closely manage the faces that appear on the matching database so that those who are clearly identified as causing a threat for a specific reason appear on it. There should be some oversight and regular re-evaluation of the persons who fall on such a list and auditable records of why they are included. Similarly, to perform the matching, the persons who appear on the CCTV footage must also be biometrically analysed. This data should be discarded as soon as it is

recognised as not a match for any one on the watch list. For ODE/AR data that shows the faces of the detected individuals does not need to be retained.

For the health data collected through the sensor data this should be stored only for the mission and if a need to retain the data for a longer period, then further permission should be sought from the data subject to ensure that they understand the reasons for the proposed longer retention period and have explicitly agreed.

The second stage of the assessment of the fundamental principles is to ensure that the necessary controls are in place to comply with the legal requirements of processing. This needs to consider the information provided to data subjects including consent, accessibility, portability, rectification and erasure, restriction, as well as identification of additional processes or potential transfers of data outside of the European Union. In the context of the last two points, it is expected that all processing would be carried out under a single controller within the deployed system and access would be limited to persons employed or authorised on behalf of the controller.

In the context of the SSO, users should be asked for their consent before providing their information and it should be freely given. Furthermore, it should be possible for users to request that their information is updated or that their account is removed if no longer needed. For the CTIP, where it is not possible to obtain consent from the data subject due to the manner and location from which the data is collected the information about the processing should be made available in some format through the data controller along with their contact details to ensure that any data subject could subsequently exercise their rights in relation to access, rectification, and erasure.

In terms of the FDR module and the ODE/AR processing, for the persons who are onsite and captured via CCTV these should be informed through notifications as they enter the site following guidance such as that from the European Data Protection Board[67] for the processing of personal data through video devices while employees onsite should also be directly informed. For the persons who appear on the matching database this should be a carefully managed process that, depending on the risk posed, may need to involve a competent authority or additional justification as to why the data subject cannot be informed.

For the data collected for the TDSS, it is expected that those having their data collected have freely given their informed consent and that they fully understand the purpose for which the data is being collected and how it will be processed. Data subjects should have the right to access any of their personal data collected during a mission for the duration of time that it remains stored within the system. They should also be able to ask for erasure of

[67] European Data Protection Supervisor (2019) Guidelines 3/2019 on processing of personal data through video devices. https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_201903_videosurveillance.pdf

such data at any point in time as well as expressing limitation on if any further data processing could take place.

4.3.3. Identification of Potential Risks and Mitigation Measures

The third phase of the DPIA is to assess the potential risks raised by the data processing activity and the overall risk level the processing poses based on the severity and likelihood of each risk occurring and the impact it would have. This should include a good understanding of the existing controls and the potential risks.

For 7SHIELD, the system is likely to be deployed, for each instantiation in a relatively controlled manner that has a limited and clearly defined scope – at least for each ground station site. Therefore, it is possible to control who is accessing the system normally, those who are entering or exiting the physical site, the locations of any CCTV cameras or the instances and areas where a UAV camera would be deployed, while those wearing the TDSS vests will be clearly identified and known to the operators. In the context of other deliverables, a thorough cybersecurity analysis of the risks and mitigation measures of the system is being carried out using the STRIDE methodology.[68] The organisational, policy and governance frameworks for which 7SHIELD would sit under would be mainly determined by the specific deploying organisation supported by 7SHIELD's comprehensive user manual that covers use and deployment of the system to support effective operation.

The SSO method itself also provides clear support for user controls through both the application of strict and role-based user controls as well as demonstrating its resistance to brute force attacks, this helps prevent unauthorised access to the wider system as well.

Privacy Risks: SSO

- Risk: Access to user's name or email address. Impact: Unauthorised exposure of contact details; increased exposure to spam email, phishing attempts; social engineering attacks for 7SHIELD and other systems where the same email address is used for login – medium; likelihood: low. Overall risk: low. Management: Use of existing security controls; requirement for non-personal email address use.

Privacy Risks: CTIP

- Risk: Unanticipated use of personal data by the data subject; lack of awareness of personal data processed within the threat intelligence platform. Impact: possible association of their personal details with cyber threats – low; likelihood: low. Overall risk: low. Management: Use of existing controls – targeted search, specific URL entry points; disposal of irrelevant information.

Privacy Risks: FDA

[68] Microsoft (2009) The STRIDE threat model. https://learn.microsoft.com/en-us/previous-versions/commerce-server/ee823878(v=cs.20)?redirectedfrom=MSDN

- Risk: Regular person matches to threat actor on matching database and triggers security alert. Impact: Distress for the citizen/visitor/employee and loss of trust in the operating organisation. Reputational risk for the organisation if reported publicly – medium; likelihood: medium. Overall risk: medium. Management: Regularly check accuracy and detection rates of models, employ human-in-the-loop to carry out manual checks on potential matches
- Risk: List of persons on the matching database is exposed. Impact: Exposure of threat information in the public domain; matching person is made aware of their presence on the database – high; likelihood: low. Overall risk: medium. Management: ensure database security, regularly evaluate list of persons on database to ensure their presence is correct.

Privacy Risks: ODE/AR

- Risk: Identifiable information is linked to specific activity which is inferred to be nefarious. Impact: loss of trust in activity recognition system – low; likelihood – low. Management: monitor for accuracy of detection algorithm.

Privacy Risks: TDSS

- Risk: Incidental finding or exposure related to health data (e.g., unusual heart rate; illness through elevated temperature). Impact: Employee is prevented from carrying out their role – high. Likelihood: low. Overall risk: low. Management: Health decisions should not made on the sole basis of information while employees should not be pressured to disclose personal medical information.
- Risk: Unacceptable use of GPS data. Impact: Access to data allows bad actor to identify the positions of the TDSS team – high. Likelihood: low. Overall risk: low. Management: Security of portable TDSS terminal

Overall, most risks appear to be in the acceptable range; however, particular attention should continue to be given to the facial recognition module (especially in view of the forthcoming AI legislation and implications for the use of biometric data) and the use of health data retrieved from the TDSS system to ensure that it only being used to support operational decisions and not the wider employment or deployment decisions.

Risks and safeguards implemented for 7SHIELD system as a whole

Additional potential risks as raised for the 7SHIELD system as a whole and previously highlighted in D9.8 include:

- Cyber-attack such as man-in-the-middle or wire sniffing
- Physical access to a datacentre
- Software vulnerability
- Password obtained through unauthorized means.

To mitigate against such potential vulnerabilities which could lead to the unauthorised access to personal data, the 7SHIELD will also develop and maintain security policies (supported by T2.3 Security Requirements) and security networking configurations. Furthermore, the provision of hardware and software security mechanisms will support the implementation of disk encryption, redundancy techniques, permission policies and role-based access control through the SSO.

5. Conclusions

This deliverable has provided an updated perspective on the legal and ethical framework for the operation of 7SHIELD first set out in D2.3. Specifically, this deliverable has ensured that due consideration has been given to updated and new legislation such as the CER Directive, NIS2 Directive and the forthcoming AI Act that has the potential to have far reaching consequences across all of Europe. Furthermore, we also carried out a data protection impact assessment that assessed the use of personal data within the 7SHIELD system and ensuring that it is adequately protected. This deliverable, in conjunction with D2.3 and the final version of the security requirements should support the final operation and future implementation of the components and the system to achieve a legal, ethical and security compliant system.

Finally, we note that both ethical norms and legislation are constantly evolving and that continued monitoring of legal and ethical compliance and the provision of appropriate safeguards is an ongoing process that should be continually monitored.