EU-HYBNET

SECOND MID-TERM REPORT ON IMPROVEMENT AND INNOVATIONS

DELIVERABLE 3.5

Lead Author: SATWAYS Ltd

Contributors: ICDS, KEMEA, L3CE, ZITiS, COMTESSA, Laurea, JRC

Deliverable classification: PUBLIC

This project has received funding from the European Union's Horizon 2020 – Research and Innovation Framework Programme, H2020-SU-SEC-2019, under grant agreement No. 883054

D3.5 SECOND MID TERM REPORT ON IMPROVEMENT AND INNOVATIONS

Deliverable number	**D3.5**	
Version:	**1.0**	
Delivery date:	**30/10/2023**	
Dissemination level:	**Public**	
Classification level:	**Public**	
Status	**File submitted for review**	
Nature:	**Report**	
Main author:	**Dr. Souzanna Sofou**	**Satways Ltd**
Contributors:	Julien Theron	EC JRC
	Michael Meisinger	ZITiS
	Alex Koniaris, Athanasios Kosmopoulos, Vanessa Papakosta	KEMEA
	Ivo Juurvee	ICDS
	Rimantas Zylius, Evaldas Bruze	L3CE
	Päivi Mattila	Laurea

DISCLAIMER

TABLE OF CONTENTS

TABLES

INTRODUCTION

I. OVERVIEW

The Deliverable D3.5 titled 'Second Mid-Term Report On Improvement And Innovations' summarises the work completed as part of Work Package 3 (WP3) titled 'Surveys to Technology, Research and Innovations' and specifically Task 3.2 'Technology and Innovations Watch', in the frame of the H2020 Project 'Empowering a Pan-European Network to counter Hybrid Threats (EU-HYBNET)'.

In more detail, the present Deliverable provides a list of Innovations and Ideas proposed to counter specific dimensions of Hybrid Threats for specific focus areas. The latter are primarily defined in the Description of Action (DoA) as 'core themes', which are studied in detail by WP2 of EU-HYBNET. The core themes represent the leading multidisciplinary methodological principles of the project, together with the Conceptual Model approach developed by the Joint Research Centre (JRC) and the European Centre of Excellence for Countering Hybrid Threats (Hybrid CoE). The EU-HYBNET project four core themes are following: 1) Future Trends of Hybrid Threats, 2) Cyber and Future Technologies, 3) Resilient Civilians, Local Level and National Administration, and 4) Information and Strategic Communication.

For every project cycle, different 'Primary Contexts/Critical Threats' are defined for the core themes. The Long list of defined gaps and needs has been presented in D2.7 by the European Centre of Excellence for Countering Hybrid Threats, while in D2.11, the JRC presented a Deeper Analysis, the short list of gaps and needs.

Based on D2.11, this Delivearble presentes ideas and innovations for each of the three primary contexts of each of the four core themes. It should be noted, that as the relevant Deliverable D2.11 is CO (consortium only), the present deliverable doesn't directly refer to the identified gaps and needs, but rather to the 'Primary Contexts/Critical Threats' these are more relevant to.

The outcomes of the present work is then provided to other Tasks of EU-HYBNET for further actions. In more detail:

i) WP3 'Surveys to Technology, Research and Innovations' / T3.1 titled 'Definition of Target Areas for Improvements and Innovations'
T3.1 will proceed with the prioritization and selection of the innovations that can be utilised by pan-European practitioners and other relevant actors to counter hybrid threats.

ii) WP2 "Gaps and Needs of European Actors against Hybrid Threats"/ T2.3 "Training and Exercises Scenario Development" and T2.4 "Training and Exercises for Needs and Gaps
D3.5 describes for T2.3 what could be the innovations that should be part of the training scenarios and eventually training activities for the most promising innovations to the identified gaps and needs.

The importance of this Deliverable for EU-HYBNET and the interactions with other Tasks is depicted in the Figure below.

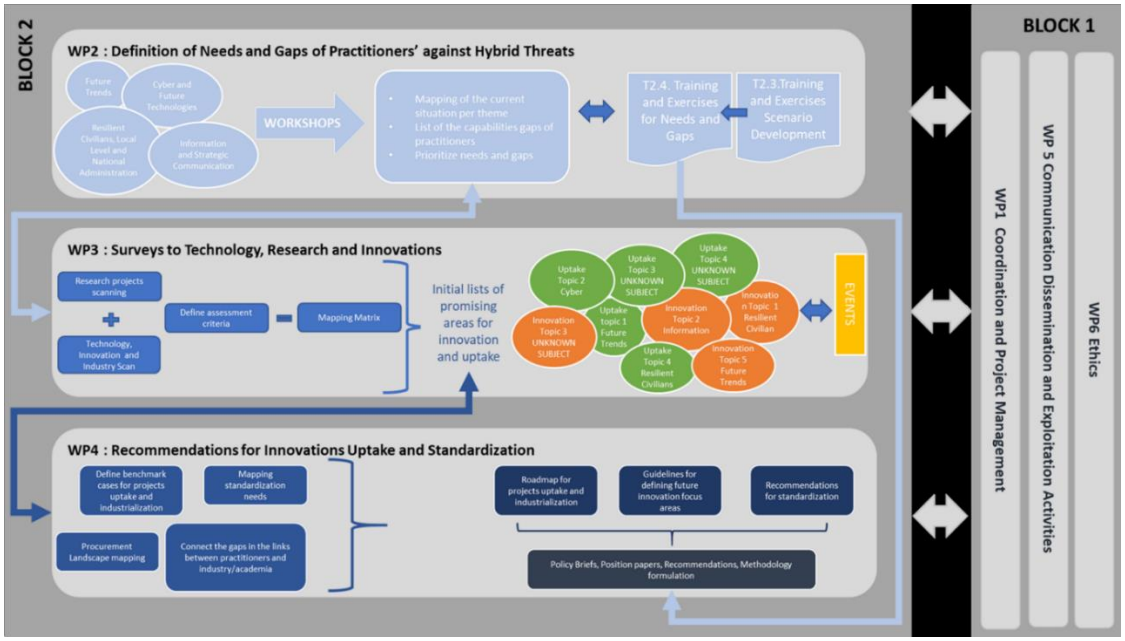

Figure 1 : EU-HYBNET structure of Work Packages and Main Activities

Additionally, D3.5 address the goals of EU-HYBNET project objective (OB) 3 and contributes to reaching its key performance indicators (KPI), as described in EU-HYBNET Description of Action (DoA) document. The OB.3 and KPI3.2 to which D3.5 delivers results are the following:

OB3: To monitor developments in research and innovation activities as applied to hybrid threats			
Goal		**KPI description**	**KPI target value**
3.2	To monitor significant developments in technology that will lead to recommending solutions for European actors' gaps and needs	Monitor existing innovations addressing gaps and needs, including areas of knowledge /performance	At least 4 reports every 18 months that address technological innovations that are able to fulfil European actors' gaps and needs

II. DEFINITIONS

Hybrid Threats

Hybrid threats aim to exploit a country's vulnerabilities and often seek to undermine fundamental democratic values and liberties[1]. Hybrid threats can be characterised as coordinated and synchronised actions by hostile state and non-state actors that deliberately target democratic vulnerabilities of governance, civic, and services spaces through a wide range of tools. The aim is to influence different forms of decision-making at institutional, local, regional and state levels to favour and/or achieve strategic goals while undermining and/or hurting the target. To effectively counter hybrid threats, improvements in information exchange, breakthroughs in relevant research, promotion of intelligence-sharing across sectors and between the EU, its MS and partners, as well as a whole-of-society policy-making approach are crucial[2].

According to the Joint Framework on Countering Hybrid Threats[1], "while definitions of hybrid threats vary and need to remain flexible to respond to their evolving nature, the Framework's conceptualisation aims to capture the mixture of coercive and subversive activity, conventional and unconventional methods (i.e. diplomatic, military, economic, technological), which can be used in a coordinated manner by state or non-state actors to achieve specific objectives, while remaining below the threshold of formally declared warfare. There is usually an emphasis on exploiting the vulnerabilities of the target and on generating ambiguity to hinder decision-making processes. Massive disinformation campaigns, using social media to control the political narrative or to radicalise, recruit and direct proxy actors can be vehicles for hybrid threats". The EU-HYBNET's definition and approach of Hybrid Threats is in line with the European Commission's report "The Landscape of Hybrid Threats. A Conceptual Model" written by the Joint Research Centre and the European Centre of Excellence for Countering Hybrid Threats (Nov 2020) [3].

Practitioners at different levels

The EU-HYBNET project follows the European Commission (EC) definition of practitioners in the security domain which states that "A practitioner is someone who is qualified or registered to practice a particular occupation, profession in the field of security or civil protection". [4] In addition[2], practitioners in the hybrid threat context are expected to have a legal mandate to plan and take measures, or to provide support to authorities countering hybrid threats. Practitioners operate at different levels of governance. Therefore, EU-HYBNET practitioners are categorised as follows: I) ministry level (administration), II) local level (cities and regions), III) support functions to ministry and local levels (incl. Europe's third sector)[2]. EU-HYBNET includes practitioner partners from all of these levels and its primary focus is on civilian security issues.

[1] Joint Communication to the European Parliament and the Council, Joint Framework on Countering Hybrid Threats, European Commission (2016).

[2] EU-HYBNET Description of Action, Coordination and Support Action, Grant Agreement No 883054.

[3] Cullen, P., Juola, C., Karagiannis, G., Kivisoo, K., Normark, M., Rácz, A., Schmid, J. and Schroefl, J., The landscape of Hybrid Threats: A Conceptual Model (Public Version), Giannopoulos, G., Smith, H. and Theocharidou, M. editor(s), EUR 30585 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-56943-5, doi:10.2760/419776, JRC123305.

[4] European Commission, MEMO - Frequently asked questions on Regulation (EU) 2019/452 establishing a framework for the screening of foreign direct investments into the Union, 09 October 2020.

Gaps and Needs

In the frame of project Work Package (WP) 2 "Gaps and Needs of European Actors against Hybrid Threats" in Task (T) 2.1 and T2.2, the EU-HYBNET project has already delivered an analysis of the pan-European practitioners and other relevant actors' gaps and needs to counter hybrid threats second time during the project. The first analysis was delivered during the EU-HYBNET 1st project working cycle (M1 – M17/ May 2020 – September 2021) and the second analysis during the second project cycle (M18 – M34/ October 2021 – February 2023). Both analyses aimed to identify, record, and understand the nature of practitioners and other relevant European actors' gaps, needs and vulnerabilities in countering hybrid threats. These include the identification of the obstacles to developing, maintaining and improving societal resilience in countering hybrid threats. D3.5 is in line with the third cycle (M35/ (March 2023 – M52/ August 2024) analysis and focus areas.

III. STRUCTURE OF THE DELIVERABLE

The document includes four main sections, each listing ideas and innovations proposed for the primary contexts of each EU-HYBNET project four core theme. The methodology applied is the same for all core themes, as detailed in the next chapter, and the structure of each section is the same. More specifically:

Section 1 addresses the first Core Theme named 'Future trends of Hybrid Threats'.

Section 2 introduces innovations for the second Core Theme named 'Cyber and Future Technologies'.

Section 3 presents ideas and innovations for the third Core Theme named 'Resilient Civilients, Local Level and Administration'.

Section 4 describes the ideas proposed for the forth Core Theme, named 'Information and Strategic Communications.

Section 5 provides the conclusions on the work performed, highlighting main focus areas and outcomes, as well as future work.

IV. Methodology

Deliverable D3.5 is the third deliverable of WP3 Task 3.2 of the EU-HYBNET project. The main actions identified at the beginning of the EU-HYBNET's 3rd project working cycle (M35/March 2023 - M52/August 2024), included:

1. A careful consideration of the results of the 1st and 2nd cycles, including the assessment of the innovations by T3.1.
2. The study of D2.7, the long list of defined gaps and needs prepared by the European Centre of Excellence for Countering Hybrid Threats, and the study of D2.11/ short list of gaps and needs written by JRC.
3. The design and preparation of a table to record, based on the long list of gaps and needs:
    - The nature of the problem, its origination, immediate outcomes, particular manifestation that needs to be addressed.
    - The need to focus on in order to solve security practitioners' and other stakeholders' Gaps, Needs and Threats.
    - The expected outcome from the use of a chosen innovation.

- The end users of this innovation.

SPACE: GOVERNANCE/CIVIC/SERVICES CORE THEME:				
What is the problem and where does it originate from? What are the parameters contributing to its escalation? Who are the pan-European security practitioners whose Gap &Need /Threat we are referring to?	What are the problem's immediate outcomes? What is the particular manifestation that we want to address? How is it threatening European Security, Democracy/ Autonomy/ Values?	What is the need to focus on in order to solve security practitioners' and other stakeholders' Gap & Need/Threat?	What is the expected outcome from the use of a chosen innovation?	Who will be the end users of this innovation? (Optional question)
To fill in one of the options below based on the spaces civic-governance-services I) ministry level (administration): II) local level (cities and regions): III) support functions to ministry and local levels (incl. Europe's third sector):				

Figure 2 : Template used for presenting Hybrid threats' related problems and expected outcomes from the use of technological solutions and innovations

4. Short descriptions of each threat were written by the T3.2 participants, that have been included as introductions in the present document.
5. A list of six scheduled discussions took place in June-July 2023 between the deliverable leader, the JRC and T3.2 partners to discuss in detail each threat, the primary contexts and the suggested ideas for countering each threat.
6. Thorough analysis of the D2.11, where the JRC presented a deeper analysis, the short list of gaps and needs.
7. Continuous monitoring of technologies and innovations, and assessment of their suitability to counter the specific dimensions of Hybrid threats.

Additionally, the dedicated template that has been designed by TNO (T3.1 leader) during the first project working cycle (M1 – M17/ May 2020 – September 2021) for presenting innovations and ideas has been updated by TNO, discussed with the Work Package leader and presented by TNO during the programmed T3.2 calls. This template helps present the innovations chosen in a coherent and systematic manner, which also enables T3.1 to proceed with assessments and comparisons of innovations in a later stage. Additionally, the use of the template supports all consortium partners in following the present work more closely and make use of the provided information. The Innovation Arena platform, developed by Laurea for the project consortium and EU-HYBNET network members, is being used throughout the entire project for this purpose.

This document suggests potential innovations and solutions to improve pan-European practitioners and other relevant actors' measures to counter hybrid threats through the analytical lens of the EU-HYBNET project four core themes.

The main innovations and ideas presented in this work are presented in the following table per core theme and primary context. An introduction to the threat is provided at the beginning of each subchapter.

Table 1 -Ideas and Innovations for Countering Hybrid Threats

CORE THEME		PRIMARY CONTEXT/ CRITICAL THREATS	IDEA/ INNOVATION PROPOSED	Partner proposing the innovation
1. FUTURE TRENDS OF HYBRID THREATS	1.1	Political Failure	Mobile application to pinpoint acts of harassment/violence on the street and online	ICDS
	1.2	New agit-prop	Anti agit-prop and hostile conspiracy warning platform	ICDS
	1.3	Alternative Reality	WeVerify, a video plugin to debunk fake videos on social media that spread conspiracy theories	KEMEA
			"EXPERIENCE" The "Extended-Personal Reality": augmented recording and transmission of virtual senses through artificial-Intelligence	KEMEA
2. CYBER AND FUTURE TECHNOLOGIES	2.1	Stealing Data/Attacking individuals	Breach Guard or Any Other Similar Available Solution	SATWAYS
			Nordplayer Or Other Similar Solution	KEMEA
			Shield, Watson Studio, Or Any Other Similar Available Solution	SATWAYS
	2.2	Online Manipulation/Attacking democracy	Code of Practice on Disinformation	L3CE
			Starlight Disinformation-Misinformation toolset	L3CE
	2.3	Attack on Services	AI And Machine Learning Technologies	ZITiS
			Advanced Surveillance Systems with Perimeter security	ZITiS
3. RESILIENT CIVILIANS, LOCAL LEVEL AND ADMINISTRATION	3.1	Spreading Violence	Expansion of the AVMS Directive	SATWAYS
			Network of anti SLAPP financial and legal support	SATWAYS
	3.2	Attack on social structure	Offline-Face-Secure-Access (OFSA)	L3CE
			Passive Authentication for Secure Identification (PASID)	
			AI-enhanced Disaster Emergency Communications	L3CE
	3.3	Undermine Inside institutions	Advanced analytical and investigative capabilities via GRACE Platform and approach	LAU
			'Antidote' to hostile messaging delivered by private messaging apps	ICDS

4. INFORMATION AND STRATEGIC COMMUNICATIONS	4.1	Media conundrum	Media Pluralism Monitor (MPM)	KEMEA
	4.2	Sectarianism	"Bad News" Prebunking Game platform	KEMEA
	4.3	Attack on Information	Real-Time Fact-Checking Browser Extension	ZITiS
			Blockchain -based verification	ZITiS

The selection of the innovations presented in the current deliverable was based on continuous monitoring of technology advances and a thorough search in various fields. The different scientific background and complementarity of the T3.2 partners allowed for a deeper understanding of possible applications of various technologies.

Additionally, during the EU-HYBNET 3rd Annual Workshop (T5.3) and 3rd Future Trends Workshop (T3.4) that took place in Bucharest, April 19th-20th 2022, the T3.2 partners had the opportunity to discuss with innovation providers-companies and consortia that were invited to participate and present their solutions in these events. Some of them were also contacted after the event in order to provide more information on the innovations they presented.

It should be highlighted that, as detailed in the Grant Agreement, the technological innovations presented aim to help European Practitioners counter hybrid threats. This deliverable also lists societal interventions, which could help practitioners protect European citizens from offensive populist influences.

1. INNOVATIONS FOR COUNTERING HYBRID THREATS: CORE THEME: FUTURE TRENDS OF HYBRID THREATS

1.1 Political Failure

EU-HYBNET responsible partner for this section: **ICDS**

1.1.1 MOBILE APPLICATION TO PINPOINT ACTS OF HARASSMENT/VIOLENCE ON THE STREET AND ONLINE

Introduction

Attacks on societal structures and cohesion, both in the form of online harassment and spread of violence, are current and there are also future trends in hybrid threats that need to be challenged. In order to be challenged successfully, the first signs of such occurrences should be noted in order to provide situational awareness if not the early warning for the attacks. In that case, the responsible law enforcement agencies can react in a timely manner in both virtual and physical space and use their resources wisely according to the situation. Should the situation still escalate, the response from rescue services may be needed as a precaution or to assist possible victims .

Involving the public has twofold purpose. Firstly, it allows to save resources of expensive online monitoring systems or CCTV and patrols on the streets. Secondly, it reduces the time needed for discovering and locating occurrences of above-mentioned threats. Thirdly, it allows to build resilience of the society itself by providing everybody – and especially the youth – an option to participate in creating more security.

BOX 1 NAME OF THE IDEA

Mobile application to pinpoint acts of harassment/violence on the street and online

DESCRIPTION OF THE IDEA

The aim of the innovation is to use readily widely available technology – i.e. smartphones – to record and geolocate acts of harassment and violence (or calls for violence) in physical space and acts of harassment and calls for violence online. Such acts my occur in the form physical action on the street, but also graffiti and or leaflets in physical space and online.

The smartphones integrate three important technologies to conduct such activity: the clock to have timestamp on the occurrence; the camera to record the action or written text as video or audio; the geolocation to pinpoint the occurrence on the virtual map. Adding option to report similar occurrences online gives the users – law enforcement agencies – the opportunity to monitor evolving situations in real time and note the correlations in physical space and online. The application would be especially useful in crisis situations like riots if used by large number of users

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  Applicable hybrid threat domains as stated by the gaps/need:
  Applicable core theme(s) as stated by the gap/need:
  The innovation contributes to controlling and preventing the spread of violence and attacks on societal cohesion. It helps to overcome lack of awareness by state institutions – especially law enforcement – and also enhancing inciting societal self-mobilization.

BOX 3 TYPE OF SOLUTION

- **Technical**
  The solution is technical in nature – by binding together existing technological solutions it opens new avenues in fight against harassment and violence.

- **Social/Human** Although technical by nature, the innovation also has strong societal and human dimension, as it provides possibility of every member of the society to participate in contributing to social cohesion and prevention of spread of hate speech. The youth would be the primary target audience to provide day-to-day content for the innovation, since they are more likely to be in places to notice occurrences – both on the street and online.

- **Organizational/Process**
  There would be a need for some organizational and technical issues in the receiving end of all information.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:** The main beneficiaries on tactical level would be law enforcement agencies that can better deal with their everyday functions in fighting occurrences harassment and violence, and especially while planning and executing their activities during crises (e.g. riots). Information gathered with the innovation would be useful also for planning further activities in the same field on ministerial level.

- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
    Information gathered by the application would provide useful data for planning further future action and resource allocation in the field.
  - **II) *local level* (cities and regions):**
    The local police and rescue authorities can use information gathered by innovation for fulfilling their day-to-day tasks, getting timely forewarning of escalation of events and informational support in crisis response.
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**
    None.
- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**The end users would be police and rescue authorities.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  TLR 8
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  The technology to be used for innovation already exists in every modern smartphone. There is no need for widespread research to be conducted.
- **Expected time to TRL-9.**
  Up to one year.
- **Expected time to market.**
  Up to two years.

BOX 6 DESCRIPTION OF USE CASE(S)

Three cases/scenarios are presented that become gradually more severe if prevention/reaction has failed in previous scenario.

Case 1: Spread of online harassment or acts of violence against some ethnic/religious/societal groups. People, especially youth, notice it online and insert the information to the application. Law enforcement agencies have, due to that early warning, higher situational awareness and may be capable of preventing escalation.

Case 2: Spread of online harassment or acts of violence against some ethnic/religious/societal groups has its implications in physical space – there emerges hate graffiti (can be directed against both – certain individuals or certain groups) and first acts of violence on the streets. Police and rescue (in the case of physical injuries or fires/arson) have higher situational awareness and may be capable of preventing escalation.

Case 3: Wide spread of online harassment or acts of violence against some ethnic/religious/societal groups has its implications in physical space and escalates to riots. Police and rescue agencies are faster and use their resources more efficiently while managing the situation.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  While assisting the fight against spread of violence and spread of disinformation online the innovation also engages the public in building resilience against attacks on societal coherence and therefore rises awareness of the public, especially youth.
- **Resilience/defensive/offensive**
  The innovation contributes to resilience of society and is defensive in nature.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  The technologies needed for/used in the innovation do exist. In addition to hardware contained in smartphones, the Geographic Information System (GIS) mapping is central.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  - Geolocation data of the smartphones can be used only with the consent of the owner of the phone, i.e. it must be allowed by the user. (If innovation is successful the settings my be applied already while selling the smartphones.)
  - The people using the mobile application for using data have to pay attention to their own security in physical space. (However, since taking pictures/videos with smartphones is widespread, using of the innovation will not be easily detectable by adversary.)
  - The innovation should not be used in physical space in the case of military action since it would pose threat to users inserting data.

<table>
<tr>
<td>

BOX 10 Implementation effort

- **Indication of costs**:

Describe the types of efforts and costs needed to implement the idea.

Innovation should be available and working in both – Android and IOS. Exact cost depends on configurations and tender conditions.

In order to be efficient, the innovation needs publicity. It can be conducted by the conducted by public affairs departments of state agencies and volunteer NGOs that limits the cost, however, some additional PR/advertisement is probably needed in order to make the use of mobile application more popular.

</td>
<td>

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?

There are no technical countermeasures available other than electronic jamming of mobile phone signals. These are mainly available for state actors, however, in the future the situation might change.
Additionally, hostile propaganda framing users of the innovation as informers for law enforcement agencies may decrease efficiency in some cases.

- **How durable is the idea (how long is the idea expected to be effective/useful?)**

The innovation is durable for the foreseeable future, i.e. it would be applicable as long as there are no major break throughs in smartphone technology or further legal restrictions.

</td>
</tr>
<tr>
<td>

BOX 12 Preconditions (optional)

- **Have all preconditions been met for the idea to be ready for implementation?**

In order to become useful the innovation needs large amount of users, i.e. there is a need for publicity.

</td>
<td>

BOX 13 Life cycle maintenance (optional)

- **Describe who will operate, maintain, update, and upgrade the described idea.**

The main beneficiaries of the law enforcement and rescue agencies, therefore, the life cycle maintenance is to be taken care of by public sector.

</td>
</tr>
<tr>
<td colspan="2">

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide

If successful, the innovation could be integrated with 112 Emergence Call System and become its integral part, thus saving considerable amount of resources.

</td>
</tr>
</table>

1.2 NEW AGIT-PROP

EU-HYBNET responsible partner for this section: **ICDS**

1.2.1 Anti agit-prop and hostile conspiracy warning platform

Introduction

The 'toolbox' of various information manipulations has grown rapidly in past years, especially as different digital and social media platforms spread. Authoritarian regimes are increasingly attempting to undermine the democratic values and rights of the EU citizens and its member states and try to polarize societies for their own strategic purposes by using various types of hybrid activities in concert with information manipulations. The design of different agit-prop and targeted conspiracy theories, which opens the way to societal polarization and radicalization are the methods of hostile information warfare and poses potential threat to the public and institutions of democratic governance.

Considering the experiences of different crises, the European Union (EU) has clearly recognized that polarizing and radicalizing disinformation, including different targeted conspiracy theories (as anti-vaccination campaigns, for instance) seriously endangers public safety and democratic values, as enshrined in the Charter of Fundamental Rights of the European Union. Some key areas of EU initiatives to strengthen the information security environment are the European Endowment for Democracy (EED), which looks to empower civil society and grassroots movements along EU borders and beyond. Second, the European Regulators Group for Audiovisual Media Services (ERGA) has been developed as a model for improving inter-governmental cooperation between the EU member states as national regulators.

The main challenge to protect democratic values and needful trust towards democratic institutions in information environment is to actively involve major media actors, especially powerful social media platforms into the coordinated process of information environment protection.

The main outcome of the innovation proposal should be better situational awareness about the spread of designed and targeted disinformation, agit-prop and conspiracy theories, more powerful analytical capabilities and better coordinated multi-layered counter-disinformation actions in social media to avoid hostile exploitation of existing political cleavages such as polarization, radicalization and undervalue of democratic institutions. It is especially important in times of mutually reinforcing poly-crises when politico-societal turbulences tend to spill-over the regions and generate negative cascading effects (in-)between different social groups and nationalities.

BOX 1 NAME OF THE IDEA

Anti agit-prop and hostile conspiracy warning platform

DESCRIPTION OF THE IDEA

The principle of the idea is based on the European Commission's Action Plan against Disinformation,[5] which states (p. 6) that 'a Rapid Alert System will be set up to provide alerts on disinformation campaigns in real-time through a dedicated technological infrastructure'. The anti agit-prop and hostile conspiracy warning platform should link the following actors:

(1) Social media users.
(2) Social media platforms and tools, such as Facebook, Telegram, X (Twitter), Signal, etc.
(3) The national SITCEN-s with EU INTCEN *via* permanent alert channels.

The social media platforms should be equipped with (special) application to monitor the content and spread of agit-prop and hostile conspiracy messages and enable users to easily (real-time) report to the platform about such content. So, the first responders could be users themselves.

The next response-layer should be the social media platforms, which also report about the "hit" to national, regional and EU level authorities who collect the incidents, analyse the trends and figure out the effective countermeasures. The rapid countermeasure could be (1) user warning about spreading malicious content, and (2) isolating further spread ('spill-over') of such content. The more advanced products could be (3) 'immunizing tools', such as short videos or quizzes with media educational effects, and (4) explanatory critical assessments for user-audiences, which are delivered in cooperation with the social media platforms.

The main outcome of the innovation proposal should be better situational awareness about the spread of designed and targeted disinformation, agit-prop and conspiracy theories, more powerful analytical capabilities and better coordinated counter-disinformation actions and campaigns in social media to avoid hostile exploitation of existing political cleavages such as polarization, radicalization and undervalue of democratic institutions.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  Improving social media users', social media platforms' and public actors' capabilities to tackle hostile information manipulations in more rapid, alert-based and coordinated manner and strengthen the protection of democracy infrastructure against hostile exploitation of political cleavages and subversive interference.

- **Applicable hybrid threat domains as stated by the gaps/need:**
  Political, Social, Informational

- **Applicable core theme(s) as stated by the gap/need:**

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

[5] JOIN(2018) 36 Final. Brussels 5.12.2018.

Better preparedness of social media users, social media platforms and relevant public actors to discover, analyse and counter agit-prop, disinformation and subversive conspiracy, which could harm the societal resilience and functioning of democratic institutions.	

BOX 4 PRACTITIONERS

Provide the applicable hybrid threat domains for which the idea is valuable: Political, Social, Informational

- **Provide the level of practitioners in the same discipline:**
  - **I) *ministry level* (administration):**
  - **II) *local level* (cities and regions):**
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Private citizens; media outlets (especially social media platforms); relevant public actors, including EU institutions

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  3
  In which stage is the solution (research, technology, available innovation, proven innovation):
  4
  Expected time to TRL-9.
  2-3 years
  Expected time to market.
  3-4 years

BOX 6 DESCRIPTION OF USE CASE(S)

The first responders could be users themselves who discover the hostile and/or subversive content and have possibility to easily report it *via* the special application. The next response-layer should be the social media platforms themselves who should report about the "hit" to national, regional and EU level POC-s/authorities who in turn collect such incidents, analyse the trends and figure out possibly effective countermeasures.

The rapid countermeasure should involve (1) rapid user warning about spreading malicious content in cooperation with social media platforms, and (2) isolating further spread ('spill-over', cross-posting, etc.) of such content.

The more advanced countermeasure products could be (3) various 'immunizing tools', such as short videos or quizzes with media educational effects to different target-audiences, and (4) explanatory critical assessments for users, which are delivered in cooperation with the social media platforms. By such multi-layered and coordinated approach, the more coherent response could be achieved.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  The innovation proposal offers advanced situational awareness between the EU institutions, its member states' governments and municipalities, empowered analytical capabilities and swiftly coordinated counter-disinformation actions in both national and EU levels to avoid hostile exploitation of existing political cleavages, especially in times of large-scale crises when political turbulences could spill-over the regions and have negative cascading effects (in-)between different nationalities and social groups, which could be targeted by hostile hybrid activities.

<table>
<tr><td colspan="2">- <u>Resilience/defensive</u>/offensive</td></tr>
<tr><td>BOX 8 ENABLING TECHNOLOGY
- Which technologies are critical in fielding the idea?
Social media monitoring tools, integrated disinformation discovery platforms.</td><td>BOX 9 Implementation
- Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?
Data and privacy protection regulations must be fully respected and technologically guaranteed.</td></tr>
<tr><td>BOX 10 Implementation effort
- Indication of costs:
Describe the types of efforts and costs needed to implement the idea.
Depends on technical configurations and tender conditions.</td><td>BOX 11 COUNTERMEASURES
Are there any potential countermeasures that could degrade the effectiveness of the solution?
Different hostile cyber-operations.

- How durable is the idea (how long is the idea expected to be effective/useful?)
3-5 years</td></tr>
<tr><td>BOX 12 Preconditions (optional)
- Have all preconditions been met for the idea to be ready for implementation?
Should be EU INTCEN (EEAS) coordinated procurement.</td><td>BOX 13 Life cycle maintenance (optional)
- Describe who will operate, maintain, update, and upgrade the described idea.
Social media platforms and EU authorities in cooperation with national and regional counter-disinformation units and private actors.</td></tr>
<tr><td colspan="2">BOX 14 MISCELLANEOUS
Any additional remarks/disclaimers/comments/information you might want to provide</td></tr>
</table>

1.3 Alternative reality

EU-HYBNET responsible partner for this section: **KEMEA**

1.3.1 WeVerify, a video plugin to debunk fake videos on social media that spread conspiracy theories

Introduction

Recent developments around the world have shown that ideological, economic, or religious polarization between contending groups is a major source of conflict and, hence, one of the key impediments to social and political progress.

The process of increasing social and political polarization promoting violence, seems to go hand in hand with economic polarization or other expressions of inequality that may inspire radicalization into physical and virtual domains.

Esteban & Ray (1994)[6] identified a set of criteria associated with a polarized society:

1.polarization is a group attribute. Isolated individuals should therefore have little weight in the calculation of social polarization.

2.within each group there should be a high degree of homogeneity.

3.there should be high degree of heterogeneity across groups.

4.the number of groups is relatively small, and each group is of significant size.

Echo chambers and isolation bubbles in the digital environment diffusing conspiracy theories, amplify the opportunities for polarized division of audiences.

While the group members show identification with each other in a polarized society, they feel socially or ideologically separated from the members of other groups. Lack of knowledge, sensitization, societal awareness of the origins of ideologies conducive of violence is outlined as a major gap.

Conspiracy theories support a process whereby people adopt extremist beliefs—including the willingness to use, encourage or facilitate violence—with the aim of promoting an ideology, political project or cause, as a means of social transformation.

In this light, it is clearly defined a need for the mapping of systemic risks caused by the spread of harmful speech on very large online platforms.

This need calls for an independent assessment and mapping of systemic risks stemming from the spread of violent, harmful, conspiracist content on very large online platforms, in keeping and complement with the stipulations of the Digital Services Act (DSA).

Countering conspiracy theories can be challenging, as they often involve deeply ingrained beliefs and emotional attachments. However, there are innovative approaches and strategies that can be effective in addressing and combating conspiracy theories:

[6] Esteban, Joan-Maria, Debraj Ray, «On the measurement of Polarization", Econometrica, Vol.62, No 4, 819-851, July 1994

BOX 1 NAME OF THE IDEA

WeVerify, a video plugin to debunk fake videos on social media that spread conspiracy theories.

DESCRIPTION OF THE IDEA

InVID WeVerify is a plugin that allows fact-checkers, journalists and any other interested users to quickly get contextual information about videos posted on Facebook, Twitter and YouTube videos.

The plugin can be found on "WeVerify", an open-source platform aiming to engage communities and citizen journalists alongside newsroom and freelance journalists for collaborative, decentralized content verification, tracking, and debunking.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  InVID WeVerify is a plugin that allows fact-checkers, journalists and any other interested users to quickly get contextual information about videos posted on Facebook, Twitter and YouTube videos. It can also perform reverse image searches on many platforms and efficiently query them. It can fragment videos, to enhance and explore key frames and images through a magnifying lens, to read video and image metadata, to check media copyright licenses, and to apply forensic filters on still images, all of which are vital tools for anyone trying to check the authenticity and source of the material.
- **Applicable hybrid threat domains as stated by the gaps/need:**
  Cyber, Information and Military /Defense domains could benefit from such solutions.
- **Applicable core theme(s) as stated by the gap/need:**
  Cyber and Future Technologies

BOX 3 TYPE OF SOLUTION

- **Technical**
  The innovation proposed is a technical one
- **Social/Human**
- **Organizational/Process**
  Whereas the innovation proposed is a technical one, it entails compliance to standards, therefore making the necessity of legal requirements and a legal framework substantial.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Cyber, Information and Military /Defense domains could benefit from such solutions.
- **Provide the level of practitioners in the same discipline:**
  The threat was listed under *Services Space* in Deliverable D2.7 (Long List of Gaps and Needs**)**
  - I) ***ministry level*** **(administration):**
  - **II) *local level* (cities and regions):**
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**
- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Media organizations, journalists, on line platforms, data hosts and providers, fact-checkers and researchers working on disinformation.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  Available innovation
  Expected time to TRL-9:
  0 years
  Expected time to market.
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

It can fragment videos, to enhance and explore key frames and images through a magnifying lens, to read video and image metadata, to check media copyright licenses, and to apply forensic filters on still images, all of which are vital tools for anyone trying to check the authenticity and source of the material.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.

It provides verification of textual claims, images, and videos (incl. AI-generated fakes); cross-modal content verification; content provenance and source trustworthiness.

Creates a decentralized database of already debunked claims and tampered images and videos, accessible both programmatically (e.g. by search engines or social platforms) and via a user-friendly web interface.

- **Resilience/defensive/offensive**
  The solution can be used in countering hybrid threats in all three manners: to promote resilience, to defend against a threat and also to provide offense against a threat.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  Pursue a holistic, cross-disciplinary approach including methods for identifying the key information sources, analysing information cascades and social network actors, and multi-modal and cross-modal content verification techniques (text, images, video).
  - The blockchain will ensure that already verified facts are recorded in an incorruptible, decentralized ledger, with no single point of failure. When the contributing verification community grows large enough, the blockchain will be absolutely necessary to ensure that individuals or groups cannot tamper with the database of known fakes to serve their own purposes.
  - New modules, released in open source, help verification professionals to identify disinformation ecosystems, the disinformation sources and the way in which disinformation campaigns are exploiting the filter bubbles, echo chambers, and highly polarized communities for maximum damage.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No

<table>
<tr>
<td>BOX 10 Implementation effort
- Indication of costs:
Describe the types of efforts and costs needed to implement the idea.
Browser Plugin free of charge</td>
<td>BOX 11 COUNTERMEASURES
- Are there any potential countermeasures that could degrade the effectiveness of the solution?
Progress in deep fake creation methodologies

- How durable is the idea (how long is the idea expected to be effective/useful?)
The solution is expected to be useful for a very long time.</td>
</tr>
<tr>
<td>BOX 12 Preconditions (optional)
- Have all preconditions been met for the idea to be ready for implementation?
No preconditions exist.</td>
<td>BOX 13 Life cycle maintenance (optional)
- Describe who will operate, maintain, update, and upgrade the described idea.
The solution is designed on a EU funded research project that provides verification systems that can help factcheckers, journalists and human rights activists to debunk and fact-check videos and images online. Many different users are using the deliverables.</td>
</tr>
<tr>
<td colspan="2">BOX 14 MISCELLANEOUS
Any additional remarks/disclaimers/comments/information you might want to provide
Keeping people safe is the number one priority for the European Union and governments across the world, and being able to respond quickly and effectively to the spread of misinformation is key. The InVID WeVerify plugin (https://weverify.eu/), funded through the EU’s Horizon 2020 research and innovation programme (https://cordis.europa.eu/project/id/825297), has already proved to be a vital tool in tackling COVID-19 related disinformation across Europe and beyond.</td>
</tr>
</table>

1.3.2 “EXPERIENCE” The “Extended-Personal Reality”: augmented recording and transmission of virtual senses through artificial-Intelligence

Introduction

Cyberspace, often associated with the World Wide Web and virtual environments, is a space where information, communication, and interaction transcend physical boundaries. It is within this boundless digital realm that the concept of alternative reality takes shape. Alternative reality refers to digital environments or experiences that deviate from our everyday reality. These realities are often crafted to provide users with unique, immersive, and sometimes fantastical experiences.

While alternative realities in cyberspace offer a world of possibilities and opportunities, they also pose significant threats that warrant careful consideration. These digital dimensions, including virtual reality (VR), augmented reality (AR), and mixed reality (MR), have the potential to disrupt our lives and societies in various ways.

The blurring of boundaries between the digital and physical realms can have profound psychological effects. Extended exposure to alternative realities may lead to a dissociation from reality, making it challenging for individuals to distinguish between what is real and what is virtual. This can result in disorientation, anxiety, and even identity crises as people become deeply engrossed in digital personas or environments.

Furthermore, alternative realities in cyberspace can be manipulated to spread misinformation and propaganda. The immersive nature of these environments can make users vulnerable to manipulation, and the spread of false narratives within these realms can have real-world consequences. It is essential to develop safeguards to prevent the misuse of alternative realities for disinformation campaigns or political manipulation.

Finally, there is a risk that alternative realities could lead to social isolation. As people spend more time in virtual worlds, there is a potential for a decline in face-to-face interactions and a weakening of social bonds. This isolation could have detrimental effects on mental health and societal cohesion.

BOX 1 NAME OF THE IDEA

"EXPERIENCE"

The "Extended-Personal Reality": augmented recording and transmission of virtual senses through artificial-Intelligence

DESCRIPTION OF THE IDEA

The EU-funded EXPERIENCE project seeks to use VR to enhance daily life by allowing brand new ways of social interaction and personal expression. It will develop the technology required to help users easily create and manipulate their own unique VR environments, significantly improving their virtual experiences. The goal is to bring VR into areas such as mental health treatment, entertainment and education, promoting VR as a means of significantly improving the human experience.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  EXPERIENCE may be exploited in many real-life contexts; for example, to evaluate the socio-psychometric characteristics (mood, character tendencies, attitudes, etc.) of a subject, highlighting any psychiatric pathologies. If the manipulation of personal virtual reality allows you to stimulate specific neuro-cardiovascular reactions, inducing particular cognitive-behavioural and emotional states in a subject, it goes without saying that the system can also be used to treat very common pathologies, such as depression, anxiety and stress.

  These include a blurred sense of reality, privacy concerns, misinformation, and social isolation.

- **Applicable hybrid threat domains as stated by the gaps/need:**

  Disinformation/Misinformation in social media environment and cyberspace in general.

- **Applicable core theme(s) as stated by the gap/need:**

BOX 3 TYPE OF SOLUTION

- **Technical**

  The innovation proposed is a technical one

- **Social/Human**

- **Organizational/Process**

  Whereas the innovation proposed is a technical one, it entails compliance to standards, therefore making the necessity of legal requirements and a legal framework substantial.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**

- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
  - **II)** ***local level*** **(cities and regions):**
  - **III)** ***support functions to ministry and local levels*** **(incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**

  Private **citizens.**.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**: **7**
- **In which stage is the solution (research, technology, available innovation, proven innovation):** Under development and research
- **Expected time to TRL-9.** 30/6/2025
- **Expected time to market.** > 30/6/2025

BOX 6 DESCRIPTION OF USE CASE(S)

This project makes real the complex interplay between multisensory perception, emotional responses, past experiences, and perspective of the future also by disentangling the mental representation of self in space and time. The new Extended-Personal Reality technological and scientific paradigms will move Europe to the future generation of extended social interactions by allowing the public at large to i) create their own VR environments as they do photos and videos without the need for technical skills, ii) create virtual simulations eliciting unique psychological, cognitive, neurophysiological, and behavioural responses, iii) automatically generate VR environments from neurophysiological data, iv) easily manipulate VR environments to communicate and elicit specific emotions.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  EXPERIENCE embeds advanced artificial intelligence routines merging information from a person's Extended-Personal Reality to inform manipulation tools including neuromodulation, multisensory biofeedback (audio, video, haptics), and subjective perception of time-space. EXPERIENCE will produce extremely realistic reproductions of the user's past and may re-administer it by modulating the associated emotional states on demand
- **Resilience/defensive/offensive**

BOX 8 ENABLING TECHNOLOGY

Which technologies are critical in fielding the idea?
Computer and information sciences, artificial intelligence, computer and information sciences, software applications, virtual reality

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No

BOX 10 Implementation effort

- **Indication of costs**:
  No indication presently

Describe the types of efforts and costs needed to implement the idea.
Data not available

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  No
- **How durable is the idea (how long is the idea expected to be effective/useful?)**
  The solution is expected to be useful for a very long time.

BOX 12 Preconditions (optional)

- **Have all preconditions been met for the idea to be ready for implementation?**
  Still under research

BOX 13 Life cycle maintenance (optional)

- **Describe who will operate, maintain, update, and upgrade the described idea.**
  The solution is designed on a EU funded research project (EXPERIENCE
  Grant agreement ID: 101017727) that provides the technology required to help users easily create and manipulate their own unique VR environments, significantly improving their virtual experiences.

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide

A plethora of innovative technological paradigms including gaming, e-learning, and neuroeconomics will be in the commercial exploitations, including the opening of a new market for actually selling EXPERIENCE.

2. INNOVATIONS FOR COUNTERING HYBRID THREATS: CORE THEME: CYBER AND FUTURE TECHNOLOGIES

2.1 STEALING DATA/ATTACKING INDIVIDUALS

EU-HYBNET responsible partner for this section: **Satways (5,10,24), KEMEA (22)**

2.1.1 BREACH GUARD OR ANY OTHER SIMILAR AVAILABLE SOLUTION

Introduction

During the last years, the exposure and tolerance to violence in the social media has unfortunately dramatically increased. As a form of online harassment, doxxing (also spelled "doxing") is a form of online harassment that means publicly exposing someone's real name, address, job, or other identifying info without a victim's consent. The aim of doxxing is to humiliate, bully, harass, or otherwise harm a victim.

This kind of attack was known for several years. An interesting article published in 2014 as a blog reveals numerous easy ways that are available to achieve this goal, in order to help internet users being more careful[7].

Recently, governments around the world have begun to pass or propose anti-doxxing laws. In the US, the state of Kentucky passed an anti-doxxing law in 2021[8], and Hong Kong passed an anti-doxxing law the same year[9]. Disclosing personal data without consent, with an intent to cause psychological harm, is now a criminal offence in Hong Kong that can be punishable by up to a HK$1 million fine and five years in jail[4]. Furthermore, doxing itself can escalate to very serious crimes, sometimes also considered doxing, such as identity theft and swatting.

Activists that work for Civil Society Organisations are often victims of such attacks and their personal and family safety is endangered. Protecting them from such attacks in a prompt and efficient way is therefore imperative.

Besides the legal framework that needs to be established, available solutions like Avast One can help protect against doxing attacks.

Additionally, the type of solution that is presented below can help prevent stealing employee data in all kinds of organisations, including politically exposed institutions. In more detail, these solutions prevent data loss, leaks, breaches and collection by third parties, which in the case of political institutions is a critical matter, as the recipients of these data are violent groups that aim to harm the individuals working in such institutions/organisations.

[7] Blechschmidt, B., Guide to doxing: Tracking identities across the web, Blog Article, November 2014.
[8] Kentucky General Assembly, 2021 Regular Session, Senate Bills, Senate Bill 267, webpage last accessed September 2023.
[9] 'What Is Doxxing, Is Doxxing Illegal, and How Do You Prevent or Report It?', Avast, Academy, online article, accessed June 2023.

BOX 1 NAME OF THE IDEA

BREACH GUARD Or Any Other Similar Available Solution

DESCRIPTION OF THE IDEA

Solutions like the Avast BreachGuard help protect personal information against data loss, data leaks, data breaches and collection by third parties, even on the dark web, and offer personal assistance by their experts whenever needed.

The solution offers monitoring for data breaches on a 24/7 basis, and protection from hackers. It automatically scans the dark web for personal information that may have been part of a data leak or data breach and helps protect the user's personal information and avoid identity theft.

Also, the solution allows reclaiming personal info from data brokers. The latter are companies that build a profile based on a user's online activity, including address, health, and financial information. Then, they can then sell this information to third parties, which can seriously impact a user's credit rating, insurance rates, and loan eligibility. Solutions like BreachGuard claim that stop these companies from collecting or selling this kind of information by removing it from their databases.

Strengthening the privacy setting of online accounts is also one of the solutions' capabilities, with the goal of reducing the amount of personal info companies have on the user, as well as stopping the spying on social media.

Such solutions may also be useful for cases of identity thefts, which are common in North America.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  The solution can be used to prevent doxing and automatically scans the web for personal information that may have been part of a leak or data breach (see description above).

- **Applicable hybrid threat domains as stated by the gaps/need:**

  Cyber, Information and Military/Defence could benefit from such solutions.

- **Applicable core theme(s) as stated by the gap/need:**
- Cyber and Future Technologies

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

  The innovation is of technical nature, but a legal framework would also be necessary in countering the problem and its routes.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**

  Cyber, Information and Military/Defence could benefit from such solutions.

- **Provide the level of practitioners in the same discipline:**

  The threat has been listed under *Civic Space* in the D2.17 Deliverable (long list of gaps and needs)
  - I) ***ministry level*** **(administration):**
  - **II)** ***local level*** **(cities and regions):**
  - **III)** *support functions to ministry and local levels* **(incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**

  Private citizens are the main beneficiaries of these kinds of technologies.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index):**
  9
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  This is a market ready product.
- **Expected time to TRL-9.**
  0 years
- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

As explained in the description, the solution can be especially useful in cases. Individuals working in journalism would especially find these technologies very useful in maintaining their anonymity.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**

  As mentioned above, these technologies can help all individuals that may be working as fact checkers and in that sense it can help societal resilience. Most importantly, activists that work for Civil Society Organisations are often victims of such attacks and their personal and family safety is endangered.

- **Resilience/defensive/offensive**
  The idea can be used in a defensive way to counter hybrid threats.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  A Windows environment is necessary. It is not known if VPN is offered, but it is also useful.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No

BOX 10 Implementation effort

- **Indication of costs:**
  Describe the types of efforts and costs needed to implement the idea.
  In the order of 60$/year

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution?**
  Cyber-attacks are constantly evolving and becoming more complex. However, cyber security as a science is also constantly developing.
- **How durable is the idea (how long is the idea expected to be effective/useful?)**
  The solution is expected to be useful for a very long time, as nothing is ever really erased from the internet, and so personal data can be stolen at any time.

BOX 12 Preconditions (optional)

- **Have all preconditions been met for the idea to be ready for implementation?**
  No preconditions are needed.

BOX 13 Life cycle maintenance (optional)

- **Describe who will operate, maintain, update, and upgrade the described idea.**
  The solution is designed mainly for Individuals. in the case where it is used by companies, the company IT department can be responsible for its maintenance.

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide
Identity theft is a crime mainly committed in North America, but it could be also committed in Europe. Such solutions can help prevent such criminal actions.

2.1.2 NORDPLAYER OR OTHER SIMILAR SOLUTION

Introduction

The leak and trading of personal and medical information from hospitals, for various purposes, such as identity theft, carries societal implications of sufficient scale. This leads to the need to place adequate consideration on the integrity of medical data.
Following the EU medical data security regulations landscape mapping, the outcomes could be tested for readiness in terms of the following aspects, which are the main things the medical community can do to strengthen their systems against a security breach:

1. Analyzing current security risks.
   Providers should conduct an annual security risk analysis for vulnerability detection and policy review, additionally making regular security audits a priority.

2. Having an incident response plan.
   Creating and implementing a response plan will help avoid escalations when a breach or incident occurs. This plan can give clear guidelines for the necessary decisions and follow-up measures.

3. Constantly educating staff.
   Employees should fully understand the consequences of a data breach in healthcare, as well as the different types of data breaches. They should also be aware of measures for both preventing a threat and dealing with one when it occurs.

4. Limiting access to health records
   It is important to identify users, track their activity, and ensure the right procedures for logging in and out of a system. Effective access permissions should be in place, depending on user position, so that only those healthcare specialists who work with medical records can access them.

5. Creating subnetworks
   Dividing a wireless network into separate subnetworks for different user groups, such as patients, visitors, personnel, and medical devices is advised. In other words, provide public wi-fi access to guests which is separate from your secure network where patient data is circulating.

6. Limiting the use of personal devices
   Healthcare professionals often use personal devices for quick remote access, but this creates additional risks, as malware entering the system makes it vulnerable to attacks. If employees are allowed to bring and use their own phones or other electronic devices for work, creation of a strict and clear policy that outlines which devices they can use within and outside the network, how to connect them to the network, and so on is crucial.

7. Avoiding the use of outdated IT infrastructure
   The older the equipment, the higher the risks of hackers accessing it. Replacing outdated devices regularly would reduce the risk of medical data breaches.

8. Updating software regularly
   Regular software updates lower the risk of cyberattacks. Having extensive expertise in healthcare software development.

9. Reviewing service-level agreements
   When choosing third-party vendors that will need access to patient data, verifying that they comply with regulations and other applicable laws should be mandatory.

10. Encrypting data
    Encryption technologies help mitigate the consequences of cyberattacks. In the USA, as per HIPAA's Breach Notification Rule, encrypted data is not considered unsecured, and so encrypted data loss does not constitute a breach.
11. Setting and enforcing retention schedules
    A retention schedule should be required. so that electronic health records (EHRs) containing sensitive data don't stay in the digital environment longer than required. This schedule should specify what information to keep, the period, storage type, and destruction methods.

12. Destroying sensitive information properly
    Confidential information should be securely destroyed.

13. Investing more in security
    Along with advanced network security tools, allocating funds to IT and legal teams is important.

    Following the above steps could help avoid data privacy breaches in healthcare, but they may not be enough to eliminate the risk of cyberattacks. Understanding more about the foundation of industry security is essential.

    Five Pillars of digital healthcare security
    Digital healthcare security is made up of five pillars: EHR systems, connected devices, payers, providers, and authorities (Government Regulators).

    1. **Electronic Health Record (HER) systems**
    2. **Connected medical devices**
    3. **Hospitals and other providers**
    4. **Health payers**
       Healthcare providers aren't the only ones who deal with personal patient records. Insurance providers, company health plans, and governmental organizations can also be targeted by attackers.
    5. **Government regulators (in the USA)**
       The medical industry is one of the highest regulated sectors. The government defines the rules for all organizations operating in the healthcare arena, including those related to the protection of patient privacy and health information:

- HIPAA is the major legislation in the field that has spurred the need for compliance in healthcare and provides standards and guidelines for handling confidential patient records.

- HITECH has tightened up enforcement of the HIPAA guidelines by implementing regular governmental audits and imposing penalties.

- The Affordable Care Act (ACA) has promoted cooperation among providers to achieve lower costs and better outcomes of care delivery by incentivizing providers who have shifted to a “pay-for-value” model instead of typical “pay-for-service.”

By implementing penalties and incentivizing compliant providers, the government can ensure that patient records are secure.

With so many factors and parties affecting data security and data quality in healthcare, the question is how risk can be mitigated.

Software solutions that help protect patient information follow:

- Cloud-based infrastructure.
- Encryption.
- Secure data standards.
- Backups.

BOX 1 NAME OF THE IDEA

NordLayer or other similar solutions.

DESCRIPTION OF THE IDEA

The specific solution has been selected for exemplary purposes only, since there are various other solutions which offer similar services.

NordLayer contributes to the protection of hospital patient data from leaking by providing secure remote access, implementing access control, encrypting data, ensuring compliance to standards in Cloud environments, providing multi-factor authentication, as well as activity monitoring and visibility.

Secure Remote Access is crucial, since healthcare organizations need modern security solutions that adapt to the complexities of today's hybrid working environments and to HIPAA rules in the USA or GDPR rules in Europe. Wherever their location, users, devices, apps, and data must have the same advanced level of network access protection.

Implementation of Access Controls is achieved by verifying all user identities before network access permissions are granted. Whoever access is granted to (enterprise users, third-party administrators, or business associates), the experience is efficient, seamless, and safe.

Data Encryption of protected health information or other sensitive data that is being sent between networks is carried out through AES 256-bit encryption, which is the most optimal solution for protecting sensitive data and minimizing cyber risks.

Ensuring Compliance to standards in Cloud Environments when using any communication service provider (CSP) such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, or others, becomes a shared responsibility between the CSP and the hospital (which is the customer). The hospital is, therefore, responsible for configuring and using cloud services in a way that complies with HIPAA or GDPR privacy requirements (USA & EU respectively).

Multi-factor Authentication (MFA) is a powerful defense against the theft of Protected Health Information (PHI) as a fundamental security measure used in many devices. NordLayer offers MFA for accessing gateways that connect hospitals and other customers to valuable resources. By following best practices in Zero Trust Network Access (ZTNA), resource access is strengthened, with the added layer of MFA protection.

Activity Monitoring & Visibility via verifying user access to resources allows businesses / organizations to understand who is inside the enterprise network. In the USA, this is one of the HIPAA requirements.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  The solution can be used to prevent the leak of hospital patient data by bridging the gap of system diversity, the outdatedness of existing systems, the inadequacy of cyber security requirements and practices, as well as the inadequacy in levels of protection and regulation of health operators.

- **Applicable hybrid threat domains as stated by the gaps/need:**

  Cyber, Information and Military /Defense domains could benefit from such solutions.

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

  Whereas the innovation proposed is a technical one, it entails compliance to standards, therefore making the necessity of legal requirements and a legal framework substantial.

- **Applicable core theme(s) as stated by the gap/need:**
  Cyber and Future Technologies

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Cyber, Information and Military /Defence domains could benefit from such solutions.

- **Provide the level of practitioners in the same discipline:**
  The threat was listed under *Services Space* in Deliverable D2.7 (Long List of Gaps and Needs).
  - I) ***ministry level* (administration):**
  - **II) *local level* (cities and regions):**
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Hospitals, healthcare providers, healthcare facilities, health insurance companies, and ultimately private citizens are the main beneficiaries of these kinds of technologies.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9
  In which stage is the solution (research, technology, available innovation, proven innovation):
  Available innovation
  Expected time to TRL-9:
  0 years
  Expected time to market.
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

NordLayer contributes to the protection of hospital patient data from leaking by providing secure remote access, implementing access control, encrypting data, ensuring compliance to standards in Cloud environments, providing multi-factor authentication, as well as activity monitoring and visibility.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.

As mentioned above, the solution can be used to prevent the leak of hospital patient data by bridging the gap of system diversity, the outdatedness of existing systems, the inadequacy of cyber security requirements and practices, as well as the inadequacy in levels of protection and regulation of health operators. The need which arises therefore is the in-depth mapping of the landscape of medical data security regulations in the EU, with the expected outcome being an enhanced benchmarking of EU hospital patient data protection actions and improved harmonization amongst EU medical practitioners. The technologies proposed by the solution can help towards compliance to these practices and they fall under the Cyber and Future Technologies theme.

- **Resilience/defensive/offensive**
  The solution can be used in countering hybrid threats in all three manners: to promote resilience, to defend against a threat and also to provide offense against a threat.

<table>
<tr><td>BOX 8 ENABLING TECHNOLOGY
- Which technologies are critical in fielding the idea?
A Microsoft Windows, macOS, Linux, Android or iOS environment are required.</td><td>BOX 9 Implementation
- Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?
No</td></tr>
<tr><td>BOX 10 Implementation effort
- Indication of costs:
Describe the types of efforts and costs needed to implement the idea.
Approximately 100€ per year (or 550€ per year including a dedicated server with a fixed IP)</td><td>BOX 11 COUNTERMEASURES
- Are there any potential countermeasures that could degrade the effectiveness of the solution?
Cyber-attacks are constantly evolving and becoming more complex. However, cyber security as a science is also constantly developing.

- How durable is the idea (how long is the idea expected to be effective/useful?)
Since patient data is stored for an unforeseeable amount of time, the solution is expected to be useful for a very long time.</td></tr>
<tr><td>BOX 12 Preconditions (optional)
- Have all preconditions been met for the idea to be ready for implementation?
No preconditions exist.</td><td>BOX 13 Life cycle maintenance (optional)
- Describe who will operate, maintain, update, and upgrade the described idea.
The solution is designed mainly for organizations, therefore, the organizations’ IT department will be responsible for its operation, maintenance, update, and necessary upgrade.</td></tr>
<tr><td colspan="2">BOX 14 MISCELLANEOUS
Any additional remarks/disclaimers/comments/information you might want to provide
As mentioned in the introduction, in the USA, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) is a federal law that required the creation of national standards to protect sensitive patient health information from being disclosed without the patient’s consent or knowledge. The US Department of Health and Human Services (HHS) issued the HIPAA Privacy Rule to implement the requirements of HIPAA. The HIPAA Security Rule protects a subset of information covered by the Privacy Rule. The Privacy Rule standards address the use and disclosure of individuals’ health information (known as protected health information or PHI) by entities subject to the Privacy Rule. These individuals and organizations are called “covered entities.”

GDPR, on the other hand, is a broader legislation that supervises any organization handling personally identifiable information of an EU or UK citizen.</td></tr>
</table>

2.1.3 SHIELD, WATSON STUDIO, OR ANY OTHER SIMILAR AVAILABLE SOLUTION

Introduction

Personal identity data is becoming massively available and transparent, therefore allowing, with the help of currently available technologies, the criminal offences of impersonation and tampering with billing systems and obligations.

In more detail, according to a recent article[10], there are 4 steps followed, that is, Discovery and investigation, Deception and hook, Attack, and Retreat. Although the tactics have not changed, phishing attacks are becoming more convincing due to the contemporary advances and capabilities of relevant technologies. The rise of deep fakes is also an alarming factor, as deception is one of the 4 stages of the attack.

Using Artificial Intelligence assisted real time fraud detection could be critical in addressing this problem, along with multistage authentication. In online fraud detection and prevention, machine learning is a collection of artificial intelligence (AI) algorithms trained with a user's historical data to suggest risk rules. By implementing the rules, certain actions can be blocked or allowed, including suspicious logins, identity theft, or fraudulent transactions. The benefits of machine learning include[11] faster and more efficient detection, reduced manual review time, better predictions with large datasets and cost-effective solution. According to a recent article[12], AI-powered fraud detection can also uncover complex and subtle patterns, thus reducing false negatives. Additionally, AI systems can adapt and evolve alongside ever-evolving fraud techniques, staying ahead of fraudsters and minimising false positives.

Another important step in online fraud detection is combining AI with biometric authentication, such as fingerprint or facial recognition. However, it should be highlighted that ethical concerns are raised by the use of AI technology in fraud detection, and these regard privacy, consent and transparency, and they should be taken into consideration.

It may also be interesting to note that AI is not only used in fraud detection but also in fraud prediction and prevention[13]. Modern collaboration platforms are using AI for that reason to stay ahead of fraudsters.

In order to support a holistic solution, related initiatives or research consortia should be identified that would collaborate with communication experts in order to communicate to the public the risks of social engineering.

[10] Shivananghan, M., Modern Engineering Explained -10 Types of Social Engineering Cyberattacks, FreeCodeCamp, March 21st, 2023.

[11] Tanant, F., Fraud Detection with Machine Learning & AI, Seon company online article, accessed September 2023.

[12] The Rise of AI-powered Fraud Detection in Payments: Securing your transactions, Sweep, May 24th, 2023.

[13] AI improves fraud detection, prediction and prevention, IBM Watson Studio, online article assessed July 2023.

BOX 1 NAME OF THE IDEA

Shield, Watson Studio, Or Any Other Similar Available Solution

DESCRIPTION OF THE IDEA

IBM Watson Studio, formerly Data Science Experience or DSX, is IBM's software platform for data science, consisting of a workspace that includes multiple collaboration and open-source tools for use in data science. In Watson Studio, a data scientist can create a project with a group of collaborators, all having access to various analytics models and using various languages (R/Python/Scala). The user can choose the tools needed to visualize or cleanse and shape data or ingest streaming data, or, most importantly to create and train machine learning models.
IBM Security QRadar Suite is a threat detection and response solution with AI capabilities, that can be used by analysts to automatically contextualise and prioritize threats. The tool can accelerate response time using advanced AI and automation and an open platform for connecting with existing legacy tools.

SHIELD comprises of three solutions.
Device Intelligence detects and stops fraud in real time with machine learning. It provides customers with knowledge regarding which users, devices and accounts are trustworthy.
AdShield blocks invalid traffic (IVT) which damages every marketing campaign and therefore revenue, and is the result of intentional and unintentional bot traffic on a website[14], whether paid or organic, data or content driven, direct or affiliate.
Compliance AI is the enterprise-grade fraud prevention comprehensive solution for the protection of the entire ecosystem. It leverages the power of device, network and artificial intelligence to profile every user, device, and account at every step of the user journey. The Hybrid AI Engine combines neural networks with symbolic AI to detect coordinated fraud attacks that other solutions miss. The user can identify suspicious patterns in real time.

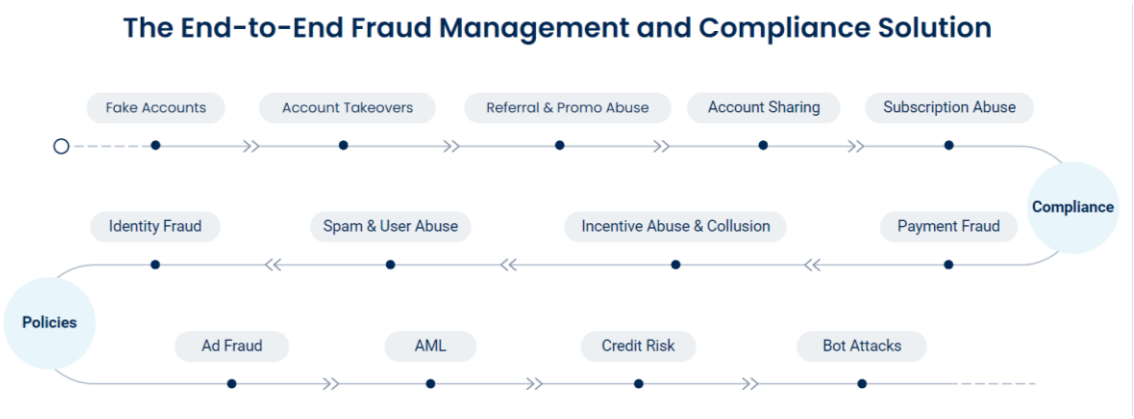

Figure 3 : Shield Compliance AI, end to end solution for fraud management and compliance.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  The solutions can be used to protect citizen's data and prevent fraud actions against them, therefore filling the gap of insufficiencies in personal data protection.
- **Applicable hybrid threat domains as stated by the gaps/need:**
  Cyber & Information domains could benefit from such solutions.

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

[14] Invalid Traffic – What Is It and How to Prevent It?, SETTUPAD blog, June 2022.

- **Applicable core theme(s) as stated by the gap/need:**
  Cyber and Future Technologies

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Economy, Cyber, Societal
- **Provide the level of practitioners in the same discipline:**
- The threat has been listed under *Services Space* in the D2.17 Deliverable (long list of gaps and needs)
  - I) ***ministry level* (administration):**
  - **II) *local level* (cities and regions):**
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**
- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Private companies

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index):**
  9
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  The suggested products are currently being sold.
- **Expected time to TRL-9.**
  0 years
- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

The SHIELD company webpage mentions various use cases, including TrueMoney (financial technology brand, providing e-payment services in Southeast Asia), inDrive (international ride-hailing service, passenger transport, operating in 48 countries), MPL (gameplay experiences)

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**

  Utilising available technologies, such as online banking and online communication, are choices of many citizens and in some cases these choices help boost the economy and social cohesion, respectively. On the contrary, the vast amount of personal data available on the internet and the ignorance with respect to risks can threaten both individuals and private companies, and jeopardise the society's well being and trust in the democratic system. It is imperative that citizens are protected and at the same time alerted of possible risks.

- **Resilience/defensive/offensive**
-

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  Typical computer operating systems

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No, there are no restrictions, the solution will be used to prevent fraud attacks.

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.
  Customised prising is offered for SHIELD based on the company characteristics. Details can be found here.

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  The threat itself is constantly evolving to overpass such technologies, but so is the solution itself.

	- **How durable is the idea (how long is the idea expected to be effective/useful?)** These technologies are constantly updated
BOX 12 Preconditions (optional) - **Have all preconditions been met for the idea to be ready for implementation?** There are no preconditions for these solutions.	**BOX 13 Life cycle maintenance (optional)** - **Describe who will operate, maintain, update, and upgrade the described idea.** The company risk analysts and data scientists will be the responsible for the solution usage and maintenance
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide** Other Fraud detection software exist that include for example Ravelin. The presented solutions are given as examples, and not as preferences. Further to the solutions identified, communication experts can drive the alerting of society about the risks of social engineering enabled by massive data availability and transparency. This can only ne accomplished with the help and guidance of the Academia. It is important also to highlight the role of legislation in order to minimize harmful threats to individuals, especially the less privileged ones.	

2.2 ONLINE MANIPULATION/ATTACKING DEMOCRACY

EU-HYBNET responsible partner for this section: **L3CE**

2.2.1 CODE OF PRACTICE ON DISINFORMATION

Introduction

Social platforms (social networks and user generated content platforms like YouTube or mass messaging like Telegram) become key instruments for public debate and interaction. In many aspects they overcome traditional media companies by impact and become default communication media in various cases.

Russia's war against Ukraine demonstrated significance of these platforms – actual information war started on the social platforms as well by bots spreading information and disinformation. Significant efforts and investments are observed by Russia trying to reach out to the western societies to impact their will to support Ukraine.

Content moderation is a common activity in all social platforms. Moderation is made to comply with legislation and to generate the traffic. Social platforms use a variety of tools to moderate content, usually this is made in the form of policies, sets of rules taxonomies embedded in technological solutions. A significant number of those policies are non-public, making it difficult to understand and initiate corrections. Those policies cover:

- Slur words policies, prohibition to use slur words (the list of slur words is not disclosed).
- Hate speech prohibition.
- Prohibition of dehumanizing speech.
- etc.

In normal circumstances these may seem like understandable policies, but in the war situation, when population within attacked and impacted countries feel existential threat, want to express their emotions against attacks on civilian population, war crimes, etc., such content moderation policies seem out of place, creating additional tensions in societies. Content moderation that becomes curtailing to freedom of expression or that follows an ideological agenda becomes a weapon that should be seriously considered. In highly emotionally charged situations (e.g. war), adversaries may leverage content moderation policies to suspend specific content and disrupt social discourse.

It is observed in practice, that Russia's activists and trolls exploit content moderation policies to suspend or remove anti-Russian content.

BOX 1 NAME OF THE IDEA

Code of Practice on Disinformation

DESCRIPTION OF THE IDEA

Disinformation is still one of the greatest risks to the European democratic information space, including components of Russia's war in Ukraine and interference in elections. As Europeans will prepare to head to polling stations in 2024, all actors must do their part in fighting online disinformation and foreign interference to protect our online debate.

Social media giants become mega force in our daily lives and manipulative aspects are highly dangerous, at the same time remains hardly reachable from regulatory perspective.

The EU came with a set of community-based tools to address the issue. So far, the best working solution is the self-regulatory Code of Practice on Disinformation. Such code of practice currently is powered by voluntary Virtual Operations Support Teams (VOST) Europe teams.

The Code of Practice on Disinformation is a first-of-its kind tool through which relevant players in the industry agreed - for the first time in 2018 - on self-regulatory standards to fight disinformation.

Its revision process was launched in June 2021 and, after the signature and presentation of the revised Code on 16 June 2022, the new Code will become part of a broader regulatory framework, in combination with the legislation on Transparency and Targeting of Political Advertising and the Digital Services Act. For signatories that are Very Large Online Platforms, the Code aims to become a mitigation measure and a Code of Conduct recognised under the co-regulatory framework of the DSA.

The strengthened Code of Practice contains 44 commitments and 128 specific measures, in the following areas.

- Demonetisation: cutting financial incentives for purveyors of disinformation
- Transparency of political advertising
- Ensuring the integrity of services
- Empowering users
- Empowering researchers
- Empowering the fact-checking community
- Transparency centre and Task-force
- Strengthened Monitoring framework

The initial concept of VOST was created in the United States in 2011, by Scott Reuter and Jeff Phillips in 2011. VOS Teams (VOST) are activated to perform specific functions in support of affected organizations & jurisdictions.

VOST Europe was founded in 2015, in a joint effort promoted by VOST Spain and VISOV, and soon after became a member of the global VOST alliance, the VOSG coalition.

Among other VOST's core missions is one relevant to hybrid threats:

> Provide support in hoax and abusive behavior, disinformation and misinformation detection, by monitoring multiple channels and by establishing direct communication channels with online platforms.

VOST Portugal leads the technological development of open source solutions that are available to other VOST in Europe. CONFIRM (CrOwd maNagement platForm for dIsaster Risk Management) is one of those tools that has been used to manage the public presence in sports events to being used as a tool to support decision making of government officials in the refugee crisis, due to the influx of refugees in Slovakia, as a result of Russia's invasion of Ukraine.

VOST Europe is also a part of the Social Media-Driven Disaster Risk Management Task Force.

The SMDRM Task Force is a consortium of researchers from universities and practitioners from disaster management organizations around the globe, facilitated by researchers from the Copernicus Emergency Services at the European Commission's Joint Research Centre.

Recently, VOST Europe have been granted mandate to deliver in practice implementation of **EU Code of Practice on Disinformation.** At this point of time the technological solution is under development.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  Covers the areas related to regulation and commitment to act for remote (foreign) privately owned social media companies, those are widely used in EU by all citizens to secure EU citizens' rights in the perspective of privacy, freedom, being informed and protected against manipulation.

- **Applicable hybrid threat domains as stated by the gaps/need:**

  Information space
  Defence
  Cyber security
  Elections
  Economy
  Journalism

- **Applicable core theme(s) as stated by the gap/need:**

  Cyber and future technologies
  Strategic Communication
  Civilian Resilience

BOX 3 TYPE OF SOLUTION

- **Technical**

  Use of generic open source tools to monitor, track and report on the activities in social media platforms.

- **Social/Human**

  Solutions is partially social – as refers to the self-regulatory type of regulation, that is rear in EU. However, currently it is the only working international regulation that enforces commitment from privately owned social media companies.

- **Organizational/Process**

  Solution is new type of collaboration between voluntary teams and public institutions.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**

  This is cross domain area and can be treated as applicable for all HT domains.

- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**

    Working with Code of Practice as regulatory.

  - **II) *local level* (cities and regions):**

    Working with VOST teams as additional tool to address implementation of Code of Practice.

  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**

    Working as bridge to the foreign social media companies.

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**

  Code of practice showed the way how it is possible to implement society driven demand for additional regulation where EU/state level regulations and directives are not working. To make it working to the full extent we need to adopt (sign) Code of Practice in all regions of EU and enforce them with local VOST teams.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  Current TRL Level of the technical solution is 7.
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  Proven Innovation but requires rollout.
- **Expected time to TRL-9.**
  1 year
- **Expected time to market.**
  2 years

BOX 6 DESCRIPTION OF USE CASE(S)

Code of Practice on Disinformation: new reports available in the Transparency Centre:
All major online platform signatories of the Code of Practice on Disinformation (Google, Meta, Microsoft and TikTok) have delivered a second set of reports on the implementation of the Code of Practice[15].

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  Major online platforms report on first six months under the new Code of Practice on Disinformation[16]:
- **Resilience/defensive/offensive**
  All areas related to social media usage.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  Self-regulatory framework.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  Not as it would be known for now.

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.
  It is running on voluntary efforts and currently do not require “named” and dedicated “for cost“ resources.

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  If other markets will come-up with similar type of instruments there will be collision of duties for the social media companies, that will put them into position to selectively follow different codes.
- **How durable is the idea (how long is the idea expected to be effective/useful?)**
  As far as it will be supported by commitment of social media platforms.

BOX 12 Preconditions (optional)

- **Have all preconditions been met for the idea to be ready for implementation?**
  It must be validated for each MS individually, as requires “self-regulatory” instruments to be legally recognized. Operations of voluntary teams of VOST must have no objector statements in the national legal basis.

BOX 13 Life cycle maintenance (optional)

- **Describe who will operate, maintain, update, and upgrade the described idea.**
  NA at the moment, but may evolve in the future once it will become standard and mandatory practice.

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide.
NA

[15] European Commission, Shaping Europe’s digital future, Major online platforms report on first six months under the new Code of Practice on Disinformation, online article, published September 2023
[16] European Commission, Shaping Europe’s digital future, Code of Practice on Disinformation : new reports available in the Transparency Centre, online article published September 2023

2.2.2 STARLIGHT DISINFORMATION-MISINFORMATION TOOLSET

Introduction

Due to huge significance of the social platforms, EU-external actors are known to invest in special instruments which leverage algorithms of platforms to promote specific content to attack democracies by manipulating online content. Such content may be targeted to justification of war, sowing distrust in democratic governments and promoting all kinds of destructive theories under disguise of free speech. Artificial amplification is one of techniques applied and it has a potential to promote this usually marginal content and expose it to wide auditorium giving the impression that this is important and legitimate “other opinion”. Artificial amplification can be used in many different circumstances (e.g.: by organised crime aiming to money laundering, etc.), but it is a clear vector of hybrid attacks, providing a necessary spread of information.

Social platforms (social networks and user generated content platforms like YouTube or mass messaging like Telegram) become key instruments for public debate and interaction. They in many aspects overcome traditional media companies by impact and become default communication media in various cases.

Russia’s war against Ukraine provides a new context of significance of these platforms – actual information wars are happening on the social platforms, bots spreading information and disinformation. Significant efforts and investments are observed by Russia trying to reach out to the western societies to impact their will to support Ukraine.

Artificial amplification can be carried out in many ways. Identification of artificial amplification is among the tasks of practitioners in hybrid threats and other fields. An example can be the elections phase where certain politicians use artificial amplification instruments to gain extensive electoral.

It is also needs to be noted that current legislation is not restricting artificial amplification in a straightforward way.

There are technological solutions for identification of bot reposts, clickbait, SPAM and other means of amplification. At the same time, social platforms seek after methodologies and tools to make them more resilient for this phenomenon.

The solution presented for this particular threat is about the complex evaluation of the content in social platforms providing access to deeper analysis of information on platforms and tools to identify artificial amplification. There are several solutions listed that can be viewed as standalone, as well as integrated comprehensive solution.

BOX 1 NAME OF THE IDEA

Starlight Disinformation-Misinformation toolset

DESCRIPTION OF THE IDEA

STARLIGHT project is one of the flagship projects dedicated to deliver easy deployable toolset to address various need of LEA and other security practitioners driven by constantly changing tech driven crimes modus operandi. In particular, STARLIGHT has one direction dedicated for disinformation and misinformation related threats. This direction is composed of several organisations developing different tooling enabling deep access of information in social platforms and tools to detect different misleading aspects of the information.

Starlight Disinformation-Misinformation toolset consist of the following toolset for version 1 and will be constantly expanding:

Tooling	Content Type	Language	Domain	Provider
Telegram Crawler	Telegram content (groups, posts, text, media)	Multi-lingual	Social networks, Discussion Forums	AIT
DeepFake detection	Is Images, Media amended	Multi-lingual	Any	AIT
Forbiden Symbol Detection	Forbidden symbolics detector	Multi-lingual	Any	AIT
Geolocalization	Recognition of recording location	Mutli-lingual	Any	AIT
Toxicity	Toxic, offensive content, comments, hateful language	German	Social Networks, Article	AIT
Story Clustering	Provides reposting chain	Multi-lingual	Social Networks, Article	AIT
Twitter Crawler	Twitter content (groups, posts, text, media)	Multi-lingual	Social Networks	AIT
Bot Detection	Is the post a bot	Multi-lingual	Social Networks	AMS
Clickbait detection	Is the post a clickbait	Multi-lingual	Social Networks	AMS
SPAM Detection	Is the post a SPAM	Multi-lingual	Social Networks	AMS
Sentiment Analysis	Provides semantic analysis of the post content based on basic emotions model	English	Social Networks, Article	AMS
Fake content Meta Detection Engine	Any post URL	Multi-lingual	Social Networks	ICCS
Toolset Integration interface	All of above	English	NA	ICCS, L3CE, AIT, AMS

There are tools dedicated to access information on general internet, communication platforms such as Telegram or X (Twitter) platforms, but majority are focused on detection of fault or forbidden content. Majority of them can work on different languages.

All of tools listed are planned to be integrated in one interface, making them easier to use.

At this point of time Starlight project is developing solutions for LEA, but it can be developed further for different target groups and serves as a good example of what is needed to handle artificial amplification complexity.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  The toolset provides possibility to analyse different communications on the social media based on its fakeness and amplification by artificial tools.
- **Applicable hybrid threat domains as stated by the gaps/need:**

BOX 3 TYPE OF SOLUTION

- **Technical**

  Toolset consisting of several tools connected by analytical purpose, based on different AI methods and techniques.
- **Social/Human**

  NA

<table>
<tr><td>For all open social media communication that plausibly was weaponized.
- Applicable core theme(s) as stated by the gap/need:
All core teams</td><td>- Organizational/Process
NA</td></tr>
<tr><td colspan="2">BOX 4 PRACTITIONERS
- Provide the applicable hybrid threat domains for which the idea is valuable:
- Cyber, Culture, Social/Societal, Public Administration, Legal, Political, Information

- Provide the level of practitioners in the same discipline:
○ I) ministry level (administration):
NA
○ II) local level (cities and regions):
LEA institutions, civil protection, hybrid units
○ III) support functions to ministry and local levels (incl. Europe’s third sector):
LEA institutions, civil protection, hybrid units

- Provide the end-users of the idea (such as NGO’s, private citizens, private companies, media outlets, police, firefighting departments):
Media outlets, police, factcheckers, hybrid threat units</td></tr>
<tr><td colspan="2">BOX 5 STATE OF THE ART
- Indication of current Technology Readiness Level (TRL 1-9 index):
TRL 8

- In which stage is the solution (research, technology, available innovation, proven innovation):
Available innovation.

- Expected time to TRL-9.
6-12 months

- Expected time to market.
1-2 years</td></tr>
<tr><td colspan="2">BOX 6 DESCRIPTION OF USE CASE(S)
Any social media content amplification of misinformation campaign detection & analysis.</td></tr>
<tr><td colspan="2">BOX 7 IMPACT ON COUNTERING HYBRID THREATS
- Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.
Solution provides toolset that gives possibility easy and semi-automated way to analyse any content or communication reliability, trustworthiness and identifies amplification techniques applied as well how much of it is being “faked” or created artificially. In addition, it provides insights about extremists, radical, criminal content usage.

- Resilience/defensive/offensive
Resilience/defensive/Investigative/Analytical</td></tr>
<tr><td>BOX 8 ENABLING TECHNOLOGY
- Which technologies are critical in fielding the idea?
AI models.</td><td>BOX 9 Implementation
- Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?
Technology is restricted access for security/LEA practitioners only during the project development.</td></tr>
</table>

<table>
<tr>
<td>

BOX 10 Implementation effort

- **Indication of costs:**
  Describe the types of efforts and costs needed to implement the idea.
- Development of relevant skillset
- Some social media API have subscription costs (i.e. Twitter)
- Solution is dockerized and do not require additional SW apart from standard dockers execution engine.

</td>
<td>

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution?**
  New models of content amplification will be introduced constantly, therefore solution must be maintained constantly and updated.

- **How durable is the idea (how long is the idea expected to be effective/useful?)**

</td>
</tr>
<tr>
<td>

BOX 12 Preconditions (optional)

- **Have all preconditions been met for the idea to be ready for implementation?**
  Users must pass accreditation for technology access, as the technology is restricted access.

</td>
<td>

BOX 13 Life cycle maintenance (optional)

- **Describe who will operate, maintain, update, and upgrade the described idea.**
  Starlight Codev01 technical team.

</td>
</tr>
<tr>
<td colspan="2">

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide.
NA

</td>
</tr>
</table>

2.3 Attack on Services

EU-HYBNET responsible partner for this section: **ZITiS**

2.3. AI AND MACHINE LEARNING TECHNOLOGIES

Introduction

Distributed Denial of Service (DDoS) attacks can seriously damage a company's reputation and business. These malicious attacks can disrupt operations, result in losses, and damage the target company's reputation.

Service disruption is one of the main effects of DDoS attacks. These attacks prevent authorized users from accessing the company's online services by flooding the target's network, servers, or infrastructure with massive amounts of traffic. Businesses can experience lost productivity and financial losses due to prolonged downtime. Businesses that rely heavily on their online presence will be particularly affected.

DDoS attacks can have a significant financial impact. Businesses can lose potential revenue or transactions when online services are disrupted for an extended period of time. Platforms for online trading, online banking and other digital businesses are particularly vulnerable to significant financial losses. For the organizations affected, such economic setbacks can have long-term repercussions.

In addition, DDoS attacks can result in customer dissatisfaction. Users who cannot access services or experience lags due to attack are likely to be irritated. This could affect the organization's ability to provide trusted services. Negative customer experiences can lead them to express their displeasure through online reviews or word of mouth, further damaging the company's reputation and turning potential customers away.

DDoS attacks are a serious problem as they damage reputation. An organization's reputation can suffer when information about an attack becomes public, especially when the incident receives significant media attention. A lack of security precautions or incompetence can be assumed if an attack cannot be effectively prevented or contained. Because of this, potential customers and business partners may have doubts about the company's reliability and credibility, which could lead to less business opportunities and cooperation.

DDoS attacks can also have legal and regulatory ramifications. These attacks may violate cybersecurity and privacy laws and regulations, depending on the jurisdiction. Businesses that fail to adequately protect their systems from such attacks, or suffer data breaches as a result, may face fines, legal action, or other penalties.

Significant investments in cybersecurity measures are required to reduce the risks of DDoS attacks. Businesses may need to invest in dedicated DDoS mitigation services, set up strong firewalls, or set up intrusion detection systems. Budgets can be tight and total cost of ownership can be impacted by these additional security costs.

In summary, DDoS attacks have a wide-ranging and severe impact on organizations. Services are disrupted, money is lost, customers are dissatisfied, brands suffer reputational damage, there are legal

ramifications, security costs are rising, and there are downsides in the market. To detect, prevent and mitigate DDoS attacks and protect their operations and reputation, organizations must prioritize the implementation of stringent security measures.

BOX 1 NAME OF THE IDEA

AI and machine learning technologies[17,18]

DESCRIPTION OF THE IDEA

Artificial intelligence (AI) and machine learning techniques play a crucial role in combating distributed denial of service (DDoS) attacks. The artificial intelligence and machine learning algorithms that form the core of this guardian's intelligence carefully monitor the ebb and flow of data.

Over time, they learn to recognize everyday patterns—the normal back and forth of traffic. If an anomaly occurs, such as a sudden increase in incoming data, alarm bells will ring. This unusual pattern is known to indicate an ongoing DDoS attack. However, these algorithms are not just limited to detecting anomalies. They are excellent for detecting anomalies that are visible to the human eye. Suspicious activity, such as a sudden rush of requests to a single server or strange packet header settings, doesn't go unnoticed. In addition, these AI-based systems are able to analyze the behavior of incoming traffic. You can spot signs of DDoS attack methods like SYN floods or UDP collision attacks. This can include redirecting traffic, blocking specific suspicious IP addresses, or applying traffic shaping policies — actions that can be critical to maintaining network integrity. It is important that these systems evolve with the threat landscape. AI-based systems can differentiate between legitimate traffic spikes and real attacks by continuously learning and improving their understanding of network behavior. This attribute is important for detecting and responding to DDoS attacks, which can result in a data deluge within seconds. Unauthorized or suspicious devices are detected and blocked from participating in the attack. In summary, artificial intelligence and machine learning are dynamic sentinels of the digital world, strengthening defenses against the relentless menace of DDoS attacks.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  Artificial intelligence and machine learning provide constant vigilance, adaptability, and accuracy, making them indispensable tools in the constant struggle to secure network infrastructure and maintain uninterrupted network services.

- **Applicable hybrid threat domains as stated by the gaps/need:**

  Industry, governments in military and commercial sectors could benefit from such solutions.

- **Applicable core theme(s) as stated by the gap/need:**

  Cyber and Future Technologies

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

The innovation is technical in nature. It also requires a process of constant improvement and adaptation to counteract new and unknown attacks.

[17] Mittal, M., Kumar, K. and Behal, S., Deep learning approaches for detecting DDoS attacks: a systematic review | SpringerLink, Soft Computing, 27, 13039–13075, 2023.

[18] Veranyurt, O., Usage of Artificial Intelligence in DOS/DDOS Attack Detection, International Journal of Basic and Clinical Studies, 8(1): 23-36, 2019.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Industry, governments in military and commercial sectors could benefit from such solutions

- **Provide the level of practitioners in the same discipline:**
  The threat has been listed under Cyber And Future Technologies in the D2.17 Deliverable (long list of gaps and needs)
  - I) ***ministry level*** **(administration):**
  - **II) *local level* (cities and regions):**
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  All are the main beneficiaries of these kinds of technologies.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  8
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  Already quite significant and growing, however ongoing improvements and adaptation needed.
- **Expected time to TRL-9.**
  1-2 years, an end of improvements will never be reached.
- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

AI and machine learning can be used in various aspects of DDoS attack prevention and mitigation across different areas like, Anomaly Detection, Traffic Classification, Rate Limiting and Traffic Shaping, Behavioral Analysis, User and Device Authentication, Dynamic Network Configuration or IoT Device Security. More or less in all network service applications.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**

- **Resilience/defensive/offensive**

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  Can be used at every platform where web services could be installed, e.g. Windows, Linux, UNIX-OSes.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.
  Hard to number this, it depends in many aspects as: Scope of Protection, Technology Stack, Data Collection and Training, Algorithm Development, Integration, Testing and Validation, Operational Training and others. It can vary from 200 k€ to x Mio €.

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  Cyber-attacks are constantly evolving and becoming more complex. However, cyber security as a science is also constantly developing.
- **How durable is the idea (how long is the idea expected to be effective/useful?)**

	The solution is expected to be useful for a very long time, since this kind of attacks will be always present.
BOX 12 Preconditions (optional) - **Have all preconditions been met for the idea to be ready for implementation?** Yes, for current state, but as mentioned before it is a changing scenario with an increasing number of requirements.	**BOX 13 Life cycle maintenance (optional)** - **Describe who will operate, maintain, update, and upgrade the described idea.**
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide**	

2.3.2 ADVANCED SURVEILLANCE SYSTEMS WITH PERIMETER SECURITY

Introduction

Physical attacks on infrastructures essential for human survival can have far-reaching and negative impacts on many sectors of society. Such attacks have severe consequences that include both short-term disruption and long-term effects.

The disruption of essential services is one of the main impacts of these attacks. Such attacks are primarily aimed at vital infrastructure such as power plants, water treatment plants, communication networks and transport systems. Power outages, water shortages, communication losses and transportation problems are therefore likely to affect the population. Such disruptions profoundly affect people's day-to-day functions, making it difficult for them to carry out important duties and maintain a sense of normalcy.

Aside from the immediate disruption, physical attacks on critical infrastructure have a significant economic impact. Sectors such as manufacturing, transport, and trade suffer setbacks due to the damage inflicted. Financial setbacks, reduced productivity and rising costs permeate the economy, threatening business, employment prospects and overall economic stability. Long-term growth and development can be hampered by consequences that can go well beyond the immediate impact.

When critical infrastructure is attacked, public safety is also at risk. For example, hospitals and emergency services could be targeted, affecting their ability to provide rapid and efficient assistance in the event of an emergency. This puts everyone's safety at risk and increases the risk to everyone's health and safety.

Physical attacks on important infrastructure often lead to social unrest and fear. Compromising access to basic needs such as food, water and health services can increase population insecurity and hardship. This can cause social unrest, civil unrest and a general sense of fear and unrest in affected communities.

Repairing and restoring infrastructure is a time-consuming and expensive process. Repairing or replacing critical systems requires critical resources, expertise and time. The affected population may continue to experience stress during the recovery period and may experience a deterioration in their quality of life.

It is important not to underestimate the psychological impact of physical attacks on vital infrastructure. These attacks can have long-term effects on individuals and communities because of the fear, trauma, and sense of vulnerability they cause. The resulting decline in mental health further exacerbates the problems for the affected population.

Governments, organizations, and communities must prioritize security precautions, backup plans, redundancy systems, and robust infrastructure maintenance strategies to mitigate the impact of such attacks and ensure critical infrastructure resilience. The continuity of critical services and the protection of critical infrastructure depend on collaboration between different stakeholders. Societies can mitigate the devastating effects of physical attacks and improve the security of their populations by investing in the protection and resilience of critical infrastructure.

BOX 1 NAME OF THE IDEA

Advanced Surveillance Systems with Perimeter security

DESCRIPTION OF THE IDEA

The implementation of advanced surveillance systems and perimeter security in the context of protecting infrastructure critical to the livelihood of the population requires the use of technologies and strategies to monitor and protect critical facilities from physical threats.

Advanced monitoring systems:
High-resolution CCTV cameras placed at strategic points within the infrastructure, focusing on critical locations where sensitive devices or necessary interfaces are used. These cameras enhanced with facial recognition and license plate recognition software to help identify and track people and vehicles entering the facility. A central monitoring station with trained personnel continuously monitors the real-time camera images transmitted via redundant channels and enables and ensures an immediate response to suspicious activities. Behavioral analysis supported by AI further improves monitoring as unusual behavior patterns are recognized and signaled. A warning is triggered when people are in restricted areas or attempt to break through the barrier. To cover areas that are difficult to monitor with fixed cameras, it is possible to use mobile surveillance units such as drones or remote-controlled cameras. Drone cameras can also be automatically sent to locations where a suspicious activity is detected.

Perimeter security:
Combined with advanced monitoring, an effective perimeter security strategy is critical. This requires intrusion detection sensors to be installed along the fence or walls. This detects unauthorized vibrations, movements or attempts to penetrate the fence. Smart fencing solutions should also be implemented that are equipped with sensors and alarms and trigger immediate warnings in the event of tampering or break-in attempts.

Integrating these perimeter security systems with advanced surveillance is critical. A detected violation automatically increases the affected area of the monitoring system. A security system configuration works to send automatic alerts to security personnel, law enforcement, and emergency responders upon detection of a breach.

To stay even one step ahead of evolving threats, continually assessing and updating these security measures is critical. Training security personnel is also important to respond effectively to security breaches and incidents. A comprehensive security strategy should include these measures as an integral part, including the development of emergency plans.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED	**BOX 3 TYPE OF SOLUTION**
- **Describe the use of the solution in reference to the gaps/need** The solutions can be used to protect against PHYSICAL ATTACKS ON INFRASTRUCTURES, closing the gap that exists in physical security protection. - **Applicable hybrid threat domains as stated by the gaps/need:** - **Applicable core theme(s) as stated by the gap/need:** - CYBER AND FUTURE TECHNOLOGIES	- **Technical** - **Social/Human** - **Organizational/Process**

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Economy, Cyber, Societal

- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
  - **II) <u>local level (cities and regions):</u>**
  - **III) <u>support functions to ministry and local levels (incl. Europe's third sector):</u>**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  Available innovation.
- **Expected time to TRL-9.**
  0 years
- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

Securing various critical infrastructures and environments such as energy facilities, water treatment, communications networks, financial institutions, government facilities, industrial complexes, transportation infrastructure, healthcare facilities, educational campuses, public events and research facilities. Security increases, preventing unauthorized access and protecting essential assets and public welfare.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  The combination of Advanced Surveillance Systems and Perimeter security contributes significantly to defending against hybrid threats through a proactive and integrated approach. It detects unconventional and diverse threats, improves situational awareness and provides robust physical defence, making it a critical part of comprehensive security strategies against hybrid threats.

- **<u>Resilience/defensive</u>/offensive**

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  Physically Infrastructure; How to install resilient

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  The threat itself is constantly evolving and therefore the techniques and solutions need to be improved and changed.

	- **How durable is the idea (how long is the idea expected to be effective/useful?)** The techniques and solutions need to be improved and changed with the evolution of threats.
BOX 12 Preconditions (optional) - **Have all preconditions been met for the idea to be ready for implementation?** There are no preconditions for these solutions.	**BOX 13 Life cycle maintenance (optional)** - **Describe who will operate, maintain, update, and upgrade the described idea.** Government, Security Contractors, Facility Owners and Operators, Security Personnel
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide**	

3. INNOVATIONS FOR COUNTERING HYBRID THREATS: CORE THEME: RESILIENT CIVILIANS, LOCAL LEVEL AND ADMINISTRATION

3.1 SPREADING VIOLENCE

EU-HYBNET responsible partner for this section: **SATWAYS**

3.1.1 EXPANSION OF THE AVMS DIRECTIVE

Introduction

Violent modes of expression, whether in speech or act, have become more accepted and mainstreamed in public discussion.
The amount of violence broadcasted by national and private television, the film industry and the social media channels has severely increased during the last years throughout Europe and the rest of the world.
The North American public's concern over the potentially harmful effects of violent television programming dates back to at least 1952[19]. Several actions have been taken ever since from the Canadian Radio-Television and telecommunications Commission[20], including the CRTC Policy on Television Violence[21].
Within the EU, audiovisual media services (including broadcasting and on-demand services) are to a broad extent regulated under the Audiovisual Media Services Directive 2010/13/EC (the AVMS Directive). In February 2014, the European Regulators Group for Audiovisual Media Services was established, which is responsible for advising on the implementation of the AVMS Directive. The AVMS Directive in particular aimed to harmonise national rules on: regulation of television broadcasts, including satellite broadcasts, under the 'country of origin', including the right for EU member states to restrict the retransmission of unsuitable broadcast content from another EU member state; promotion, production and distribution of television programmes within the EU, including quotas for European-produced content and content made by independent producers; access by the public to major (sports) events; television advertising, product placement and programme sponsorship; **protection of minors from unsuitable content**; and right of reply (of any natural or legal person whose legitimate interest has been damaged by an assertion in a television programme).
Expansion of 2010/13/EC (the AVMS Directive) with stricter rules on protection of minors from unsuitable content would probably be a solution in the value-added direction.
For this threat, a technological innovation was not found, but the analysis was included in the deliverable for the sake of completeness.

[19] Alter, S., Violence on television, Law and Government Division, publications of the Government of Canada, October 1997.

[20] Canadian Radio-Television and Telecommunications Commission, Government of Canada, assessed July 2023.

[21] CRTC Policy on Violence in Television Programming, 1996-36, Canadian Radio-Television and Telecommunications Commission webpage, assessed July 2023.

3.1.2 NETWORK OF ANTI SLAPP FINANCIAL AND LEGAL SUPPORT

Introduction

During the last years, SLAPP attacks (Strategic Lawsuits against Public Participation) have become more intense, thereby threatening democratic values and rights. According to the Commission,[22] SLAPPs are manifestly unfounded or abusive court proceedings. They are a particular form of harassment increasingly used against journalists, human rights defenders and others engaged in public participation in a matter of public interest and upholding democratic values and fundamental rights.
The Expert Group Against SLAPP (E03746, JUST- DG Justice and Consumers) was founded in March 2021, with the mission of advising the Commission on any matter relating to the fight against SLAPP or the support to their victims.
Among their members are A) individual experts, appointed in his/her own personal capacity, B) or as representative of a common interest, C) organizations like the Council of Bars and Law Societies in Europe (CCBE), who have also written a policy paper CCBE Position on abusive litigations targeting journalists and right defenders, the European Federation of Journalists (EFJ), who claimed that the Commission adopted a watered-down position on anti-SLAPP directive, the News Media Europe (NME), who also published a position paper on the proposed EU Directive anti SLAPPs , and other public entities.
It is interesting to highlight that one of the topics discussed in the agenda of one the Expert Group meetings [23] (21-11-2022) was the funding opportunities to buttress organizations that provide guidance and support for such targets. In particular, the CERV Programme (Citizens, Equality, Rights and Values) has been mentioned[1], which aims to promote rights and EU values by providing financial support and capacity building for civil society organisations, including those active in anti-SLAPP practices. In fact, protection from SLAPP falls under the EU Charter of fundamental rights call.
Additionally, the Creative Europe Program aims, among other objectives, to provide legal and practical support (sheltering program and financial support) to journalists and other media practitioners in need, including targets of SLAPP.
Besides the legislative initiatives that the Commission is leading, it is important to ensure that an ongoing investigation will not be stopped because of a lawsuit. A potential countermeasure would be to identify or create a consortium of journalists that would support and enable each other's investigations when some of them would be silenced by strategic lawsuits and abusive litigation.

[22] European Commission, The 2023 CHAR-LITI Call for proposals under the CERV, January 26th 2023.
[23] European Commission Expert Group against SLAPP, Minutes of Meeting, 21 November 2022.

BOX 1 NAME OF THE IDEA

NETWORK OF ANTI-SLAPP FINANCIAL AND LEGAL SUPPORT

DESCRIPTION OF THE IDEA

The European Anti-SLAPP Conference (third edition to take place in London and online on November 27th-28th 2023) will focus on tracking implementations of SLAPP solutions. It is organized by The Foreign Policy Centre (FPC, an outward-looking, non-partisan international affairs think tank based in the UK), the Justice for Journalists Foundation (JFJ, a London-based charity whose mission is to fight impunity for attacks against media) and the International Bar Association's Human Rights Institute (IBAHRI, which works with the global legal community to promote and protect human rights and the independence of the legal profession worldwide). The conference is also supported by a list of other organisations.

The conference web page also presents resources, that is, a list of initiatives to address SLAPP as well as ways to acquire practical and legal support.

Firstly, a dedicated tool (offered in 5 languages) helps journalists understand if they are facing SLAPP.

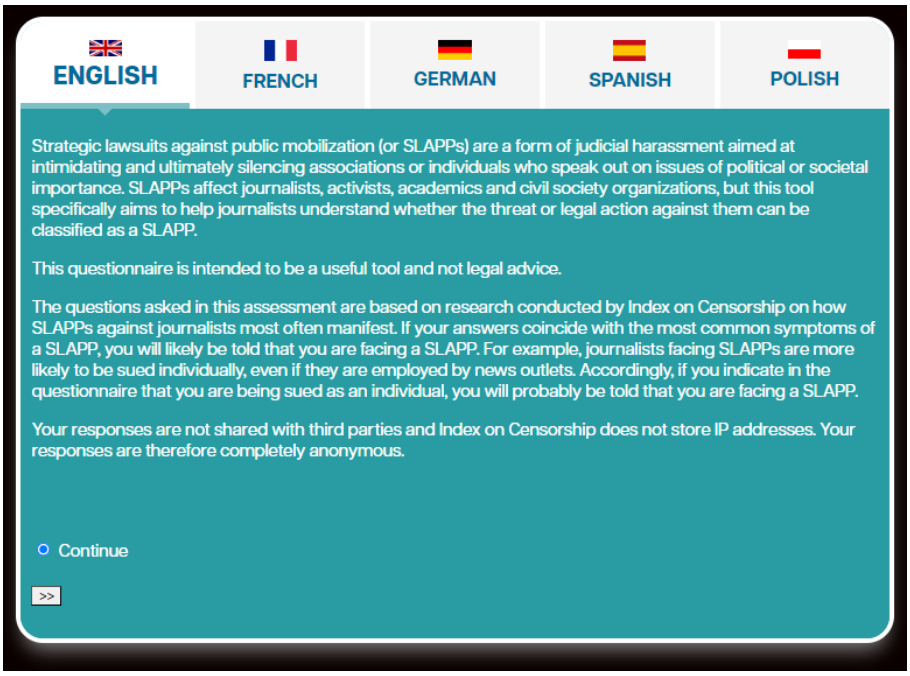

Figure 4 : Screenshot of a tool offered by Index on Censorship for helping journalists understand wheather the threat or legal action against them can be classified as a SLAPP

Then, means to acquire legal support are presented, and these include:

the European Centre for Press and Media Freedom (ECPMF), that proffers and coordinates legal support on matters related to free speech for individuals and organisations working in countries located geographically in Europe. Depending on ECPMF assessment, such support may consist of Financial support to cover lawyer's fees, General guidance, Access to expertise in policy and law making, Engagement in national or international litigation Provision of independent analysis, observation or advocacy around a case.

Furthermore, the Coalition against SLAPPs in Europe (CASE) provides a map based directory of law firms providing pro-bono legal support across Europe.

Finally, there are options for reporting the case, and these include CASE and if the particular case pertains to the UK, the UK anti-SLAPP coalition.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED	BOX 3 TYPE OF SOLUTION
- **Describe the use of the solution in reference to the gaps/need** The solution is meant to be used as a response to the vulnerable situation of journalists in light of SLAPP. - **Applicable hybrid threat domains as stated by the gaps/need:** - **Applicable core theme(s) as stated by the gap/need:** Resilient Civilians, local level ad administration	- **Technical** - **Social/Human** - **Organizational/Process.**

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
  - **II)** ***local level*** **(cities and regions):**
  - **III) support functions to ministry and local levels (incl. Europe's third sector):**
- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Journalists facing SLAPP

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  The network is available and active.
- **Expected time to TRL-9.**
  N/A
- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

Use cases can be found via the relative webpages

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  The network can be used to stop the intimidation of journalists which forces them to spend an enormous amount of money and energy and prevents them from doing their service to democracy, which has a profound impact on media freedom.
- **Resilience/defensive/offensive**

BOX 8 ENABLING TECHNOLOGY	BOX 9 Implementation
- **Which technologies are critical in fielding the idea?** No technologies needed other to safe access to the internet.	- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?** No
BOX 10 Implementation effort	**BOX 11 COUNTERMEASURES**
- **Indication of costs**: **Describe the types of efforts and costs needed to implement the idea.** There are no costs , the solution is meant to support the journalists legally and financially.	- **Are there any potential countermeasures that could degrade the effectiveness of the solution**? **How durable is the idea (how long is the idea expected to be effective/useful?).** The idea has a great potential to stay effective; in fact, the network is only expected to grow and become stronger.
BOX 12 Preconditions (optional)	**BOX 13 Life cycle maintenance (optional)**
- **Have all preconditions been met for the idea to be ready for implementation?** No preconditions in this case.	- **Describe who will operate, maintain, update, and upgrade the described idea.** The organisations listed above regulate their operation.
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide**	

3.2 ATTACK ON SOCIAL STRUCTURE

EU-HYBNET responsible partner for this section: **L3CE**

3.2.1 OFFLINE-FACE-SECURE-ACCESS (OFSA)

Introduction

Academic institutions, especially higher education institutions, by their nature are usually open, mutual trust based and manifesting freedom. Security concerns, especially security of new knowledge created in these institutions, are not of the highest priority. Comprehensive culture of security not conflicting with academic openness and freedom should be sustained while findings ways to secure critical information and research capital. On the other hand, there is plenty of evidence, proving that adversaries are exploiting this vulnerability by different means, from making use of students to more sophisticated espionage techniques.

The spectrum of security building blocks is very wide. The starting point can be the definition of critical resources to be specifically handled. There is a good example of good practice in the USA. The "Critical and Emerging Technologies List Update"[24] defines critical and emerging technology areas that are of particular importance to the national security. There are 19 different areas listed, each described in subfields. The Update is related to "The National Strategy for Critical and Emerging Technologies" (2020) and "Interim National Security Strategic Guidance" (2021). The List is used, among others, to protect sensitive technology from misappropriation and misuse. Research and other assets related to the listed technologies should be secured and respective liabilities are related to them. The same can be defined at EU or national levels applying regulatory measures. At this point we are not suggesting any specific regulations, this goes more as a proposal for policy makers as this deliverable is focused on technological solutions.

Further higher education institutions can explore clearance and access management methodologies used by more security concerned institutions and to apply them to an acceptable level for researchers, administration, teachers or even students. There are different solutions for digital access management, download controls and file sharing tracking. It can go further to the mapping of political attitudes and identification of potential treat points. Those are just a few to think of. Main question remains open: how we can induce security culture without endangering value system and security being not in focus of activities in such institutions.

To avoid exclusion of certain groups (e.g. applying some changes only for students or administration, or only for third country researchers, etc.) and to make it easy to handle and accept, physical access management systems can be applied. They can also be combined with the access of digital assets. If handled properly they do not object to the nature of institutions and can be one of the entry points for security culture moulding. There are many access management systems in the market. In this document two innovative solutions are presented that might suite needs of higher education

[24] Critical and Emerging Technologies List Update , Executive Office of the President of the United States, a report by the Fast Track Action Subcommittee on Ctitical and Emeging Technologies of the National Science and Technology Council, February 2022.

institutions. Those serves as examples of how very sensitive and high importance national or EU security assets can be better protected. This should not be applied to the overall activities of education institutions, just to very specific sections and activities.

Access management solutions can make impact on several aspects, like physical access of unwanted visitors, divide infrastructure into clear security zones, raise awareness of security importance, etc. It is one of the security building blocks and many relevant aspects remain unsolved, but it can be a good starting point.

BOX 1 **NAME OF THE IDEA:**

Offline-Face-Secure-Access- (OFSA)

DESCRIPTION OF THE IDEA

To better protect employees and assets in sensitive areas, systems must be based on a high level of security, and innovation to be efficient.

OFSA prototype is constructed to provide an innovative and highly secure offline access control solution to dedicated and sensitive areas.

The first goal of this prototype is to create the first autonomous double offline authentication for access control that will be highly secured and more efficient against cyber threats. The second goal is to create a face biometric recognition for access control with no data stored on the solution that will be GDPR compliant.

The solution is developed along three predefined axes:

1. AK1 Minibox validates that anyone in a protected zone is allowed to be in this zone simply with his smartphone by an authentication of his phone number.
2. Biometric authentication Face Alive by ID3, to detect real humans' and authenticate their identity.
3. In Offline mode, unification, and deep level integration of communication between access control and Biometric Face Identification technology to validate the access and open any electrified opening system.

Validation process before using the solution (must be done one time only):

1. Administrators use Akidaia solution to create access rights to the mobile app of the user.
2. The user goes to the ID3 /Face Alive Enrolment to Create a QR code with face biometric specifications of the user, send it to the AK1 minibox and delete it from the control kiosk. The AK1 minibox sends it to the phone of the user and deletes it from the AK1 minibox. The phone of the User will be the only one to store this QR code.

Validation upon use:

1. The user comes to the control kiosk facing the camera.
2. Communication between the Akidaia mobile application, the AK1 Minibox and the control kiosk.

If the Biometric specifications and the Cryptographic challenge between the three bricks are correct the electrified opening system will open.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

Describe the use of the solution in reference to the gaps/need

The solution can be employed to safeguard higher educational institutions from external interference, such as the penetration of sensitive facilities and highly confidential projects by foreign powers. The solution is specifically designed to enhance the level of security of critical information (e.g., access control of unwanted persons), when employed in tandem with a robust collaboration with law enforcement agencies (LEA) and intelligence services, it presents a promising foundation for fortifying the security of social structures.

Applicable hybrid threat domains as stated by the gaps/need:

Political, culture, social/societal, administrative, and informative domains.

Applicable core theme(s) as stated by the gap/need:

Core Theme 1. Resilient civilians, local level, and administration.

BOX 3 TYPE OF SOLUTION

Technical

Social/Human

Organizational/Process

The primary focus of this innovation lies in its technical aspects. However, it is imperative to deliberate on the legal and ethical implications before deploying this solution in operational use.

BOX 4 PRACTITIONERS

Provide the applicable hybrid threat domains for which the idea is valuable: Cyber, Information, Intelligence, administration can benefit from the proposed solution.

Provide the level of practitioners in the same discipline:

- *ministry level (administration).*
- *local level (cities and regions)*: Higher education institutions, programs, teachers, professors, students.
- *support functions to ministry and local levels (incl. Europe's third sector).*

Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):

This innovation has a potential to become highly beneficial for any sensitive infrastructure operators.

BOX 5 STATE OF THE ART

Indication of current Technology Readiness Level (TRL 1-9 index):

The Akidaia company was created in 2020 and reached the TRL9 in the beginning of 2023. The offline identification technology is now ready for sale with an international patent pending solution (2021). ID3's facial recognition technologies have already reached TRL 9 and are sold as licenses, software, and web services to enterprise customers. Face-Alive's antispoofing have already reached TRL 9 with its ISO

30107-3 compliance certification. The biometric library developed by ID3 integrated into the JCOP ID is NXP's first solution offering biometric match-on-card with CC EAL6+ certification, which enables highly secure biometric based user authentication (2022). Innovation in high quality and secured biometric acquisition, is validated by 2 patents held by KIS partner, PCT/EP2021/072042 -2021 (2021).
The combined technologies will be fast, reliable and don't require any complex challenge for end users.

In which stage is the solution (research, technology, available innovation, proven innovation):
The technologies are available; however, the integration of different components might take an additional 12 months.
Expected time to TRL-9: 1 year.
Expected time to market: 1 year.

BOX 6 DESCRIPTION OF USE CASE(S)

The solution encompasses a fully isolated biometric access control system, wherein the user retains exclusive ownership of their biometric data. This approach offers a distinct advantage when compared to solutions that rely on centralized databases or online biometric processes. Presently, the Asian market exhibits notable dynamism, and providers are increasingly recognizing the significant opportunities within the European market. They are making substantial investments in access control technology; however, a key challenge they face is aligning their offerings with European legislation, particularly concerning the protection of personal data.
The solution can be employed in government buildings and institutions ensuring that access control remains robust and secure. Manufacturing units that often have restricted areas where only authorized personnel should have access, power stations and utility infrastructure as critical assets, higher educational institutions to protect facilities and sensitive data.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs:
OFSA solution contributes to countering hybrid threats by providing a multi-faceted security approach that combines physical security, sensitive data protection, and deterrence. It strengthens the resilience of educational institutions against a high range of threats, including those that involve a combination of physical and cyber elements. The solution contributes to critical gap named Insufficient protection of social structures to potential attacks and critical need: Correction of the social structures' fragility.
Resilience/defensive/offensive:
The technology can be applied to enhance the resilience of higher educational institutions in the face of hybrid threats.

BOX 8 ENABLING TECHNOLOGY

Which technologies are critical in fielding the idea?

BOX 9 Implementation

ID3's facial recognition technologies. Face-Alive's anti-spoofing technology. The biometric library developed integrated into the JCOP ID is NXP's first solution offering biometric match-on-card with CC EAL6+ certification, which enables highly secure biometric based user authentication (2022) Innovation in high quality and secured biometric acquisition, is validated by 2 patents held by KIS partner, PCT/EP2021/072042 - 2021 (2021). The combined technologies are fast, reliable and don't require any complex challenge for end users.	**Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?** Considering that technology includes biometric data components, it is imperative to ensure compliance with GDPR regulations EU and national legislation.
BOX 10 Implementation effort **Indication of costs. Describe the types of efforts and costs needed to implement the idea:** The prototype is developed by the consortium of 5 SMEs and here is the role of each SMEs: • **Akidaia- SME-France:** Offline access control using cryptographic challenge authentication with smartphones, • **ID3 - SME-France:** Provide biometric algorithms in the fields of facial recognition, • **Aldo Ferraro orafo Jewellery - SME - Italy:** make operational tests on the becoming prototype (LOI), • **Fotofix GMBH - NotSME - Germany:** make opérational tests on the becoming prototype (LOI), • **KIS- NotSME - France:** Develop kiosk hardware and soft to meet biometric performance and requested form factor. The total budget for different components integration: 60K EUR	**BOX 11 COUNTERMEASURES** **Are there any potential countermeasures that could degrade the effectiveness of the solution?** Comprehensive and advanced security measures must be carefully considered and subsequently implemented to safeguard educational institutions from foreign powers' attacks. OFSA solution is the combination of different technologies that are designed to ensure the minimum possibilities to degrade the effectiveness of the solution. However, the foreign powers may employ a wide range of cover and overt tactics to compromise the educational institutions. These measures encompass not only physical security but also address cyber threats, social engineering, and other methods that foreign power may employ. By implementing complex security measures including close collaboration with LEA and intelligence services educational institutions can significantly enhance their resilience and safeguard their core mission of providing education, research and knowledge dissemination. **How durable is the idea (how long is the idea expected to be effective/useful?)** Akidaia has created a 100% offline identification solution. The innovative character of this technology has been validated by the registration

	of a patent under the N° FR2113335 in December 2021, which received favourable opinion for France and Europe. The big innovation is the fact that in offline mode, during this double cryptographic and face recognition, no data will be stored in any of the two devices, each demand will be unique and deleted directly in the two devices (AK minibox and the ID3- Face alive camera, finally the only place where the Biometric data will be stored is the phone of the user. The innovation represents a ground-breaking advancement in the market and its potential towards gaining market share is currently underway. This innovation just started to capture the attention of different stakeholders and has the potential to shape the future of its respective industry.
BOX 12 Preconditions (optional) **Have all preconditions been met for the idea to be ready for implementation?** All preconditions are met. A consortium of 5 SMEs has joined forces with the goal of creating a product that is fully prepared for the market. This partnership involves combining their resources, expertise, and technologies to create a product that meets market demands and make it ready for commercialization.	**BOX 13 Life cycle maintenance (optional)** **Describe who will operate, maintain, update, and upgrade the described idea.** Consortium of 5 SMEs will take responsibility for the maintenance and support of OFSA solution.
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide:**	

BOX 1 NAME OF THE IDEA:

Passive Authentication for Secure Identification- PASID

DESCRIPTION OF THE IDEA

Sensitive infrastructures, whether physical or digital, rely on active authentication systems that use credentials such as usernames, passwords, PINs, access cards, or biometric scans to verify a user's authorization. However, these systems can be problematic, as they only check the credentials provided by the user, not the individual providing them. This can lead to unauthorized access when someone copies a user's credentials.

PASID is password-less system that continuously authenticates individuals in a frictionless, privacy-preserving manner. It is based on behavioural analytics engines that utilize machine learning to transform the user's activity data into behavioural biometrics. The system includes a companion smartphone app and a cloud platform. The app records the user's activity data through accelerometers, gyroscopes, and other sensors and securely sends it to the cloud for processing by behavioural analytics engines. The resulting behavioural biometrics allow the system to authenticate users continuously without requiring extra input. The system authenticates the human, not just the provided credentials, enabling robust digital security and a superior user experience all the time. PASSID solution can be integrated through a simple API, similar to any other multi-factor authentication (MFA) service the educational institution uses. After users install the companion application, the system can start authenticating them immediately. When a user requests access to sensitive infrastructure, the system verifies and authenticates the individual in addition to the traditional credentials requested, stopping unauthorized individuals from using stolen credentials.

The system can continue to authenticate a user's journey through the infrastructure to provide additional security. Thanks to the ability to continuously authenticate individuals, the system can also authenticate user data such as steps taken, health statistics, and more.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

Describe the use of the solution in reference to the gaps/need:

The solution can be employed to safeguard higher educational institutions from external interference, such as the penetration of sensitive facilities and highly confidential projects by unwanted persons. Furthermore, the PASID solution will protect the sensitive infrastructure beyond the point of entry as it will continue monitoring the user's activity

BOX 3 TYPE OF SOLUTION

Technical
Social/Human
Organizational/Process

The primary focus of this innovation lies in its technical aspects. However, it is imperative to deliberate on the legal and ethical implications associated with the collection of biometric data prior implementation of PASID solution in operational contexts.

throughout his journey inside the sensitive infrastructure. **Applicable hybrid threat domains as stated by the gaps/need:** Political, culture, social/societal, administrative, and informative domains. **Applicable core theme(s) as stated by the gap/need:** Core Theme 1. Resilient civilians, local level, and administration.	

BOX 4 PRACTITIONERS

Provide the applicable hybrid threat domains for which the idea is valuable: Cyber, Information, Intelligence, administration can benefit from the proposed solution.
Provide the level of practitioners in the same discipline:

- ***ministry level (administration),***
- ***local level (cities and regions):*** Higher education institutions, programs, teachers, professors, students.
- ***support functions to ministry and local levels (incl. Europe's third sector).***

Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments): This innovation has a potential to become highly beneficial for sensitive infrastructure operators, private operators, educational institutions, LEA.

BOX 5 STATE OF THE ART

Indication of current Technology Readiness Level (TRL 1-9 index): Current TRL-5.
In which stage is the solution (research, technology, available innovation, proven innovation): Piloting in real environment.
Expected time to TRL-9: 2 years.
Expected time to market: 2 years.

BOX 6 DESCRIPTION OF USE CASE(S)

The strategic roadmap of technology provider begins with the implementation of gait-based behavioural engine focusing on protecting sensitive infrastructures that require physical access. Subsequently, their expansion plan encompasses the incorporation of additional behavioural engines and penetration into other markets.

In the initial plan to prove the technology in Europe the technology providers are planning to engage users in large companies with significant sensitive infrastructure. In the first step, they will work with the companies to identify subgroups of their end-users and ask them to install the companion app on their devices while integrating the API into their authentication system. This will enable them to gather feedback from the end users and the companies about their experience with the system so that it can be tailored to their needs and ensure that it is easy to use.

The ultimate goal is to work closely with these companies to ensure that PASID system meets their needs and is easy to use while also providing a highly secure and reliable multi-factor authentication service. By taking a deliberate and thorough approach to testing and implementation, they can help these companies enhance their security posture and protect their sensitive data from unauthorized access.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs:

The novel security approaches in conjunction with a close collaboration with security agencies strengthens the resilience of educational institutions against a broad spectrum of threats, including those that blend both physical and cyber elements.

PASSID solution contributes to critical gap: Insufficient protection of social structures to potential attacks and critical need: Correction of the social structures' fragility.

Resilience/defensive/offensive:

The technology can be applied to enhance the resilience of social structures in the face of hybrid threats.

BOX 8 ENABLING TECHNOLOGY

Which technologies are critical in fielding the idea?

1. Use state-of-the-art machine learning algorithms and innovative architecture to capture and analyse

BOX 9 Implementation

Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?

the behaviour underlying activity data through the smartphone's accelerometers, gyros, and other sensors. 2. Introduced a companion app and an API instead of an SDK to protect all digital assets immediately and avoid needing app-specific training. 3. System protects not only the sensitive infrastructure at the point of entry but also the smartphone device. It would understand instantly if an unauthorized person took over a smartphone connected to the sensitive infrastructure and alert the security team. 4. System will protect the sensitive infrastructure beyond the point of entry as it will continue monitoring the user's activity throughout his journey inside the sensitive infrastructure. Thanks to the ability to continuously authenticate individuals, the platform can also authenticate user data such as steps taken, health statistics, and more.	The consortium claims that as the data is under the direct management of the user, there are no data privacy implications.

BOX 10 Implementation effort

Indication of costs. Describe the types of efforts and costs needed to implement the idea:

The consortium is comprised of 4 SMEs: 3 from Greece, and 1 from Estonia. The Greek companies will focus on technology and business development in South-eastern Europe. They will leverage Greece's strong technical skills, which are available at reasonable prices. One of the companies, thanks to its strong business network, will pursue proof of concepts. The Estonian company will focus on business development in the Nordics and Northern Europe, taking advantage of Estonia's high profile in the field of cybersecurity within the European community.

The goal of solution provider is to create a subscription-based business model that provides customers with access to PASID platform on a software-as-a-service basis. The subscription will be based on the number of users, and the level of interaction and API calls one company is expected to have. This approach will enable to provide the customers with a flexible and affordable pricing structure that meets their needs.

BOX 11 COUNTERMEASURES

Are there any potential countermeasures that could degrade the effectiveness of the solution?

PASSID solution is the combination of mature technologies that are designed to ensure the minimum possibilities to degrade the effectiveness of the solution.

However, the foreign powers may employ a wide range of cover and overt tactics to compromise the educational institutions. These measures encompass not only physical security but also address cyber threats, social engineering, and other methods that foreign power may employ. By implementing complex security measures including close collaboration with LEA and intelligence services educational institutions can significantly enhance their resilience and safeguard their core mission of providing education, research, and knowledge dissemination.

How durable is the idea (how long is the idea expected to be effective/useful)?

How durable is the idea?

The idea is durable by the consortium of 4 SMEs the team has diverse skills and experiences in business and technologies development, e.g. the machine learning and data engineering experts on the team will be responsible for developing the behavioural engines that are at the core of passive authentication technology. The full-stack developers will be responsible for building the back end and API of the system. The front-end developers will be responsible for developing the companion app that collects activity data from the user's device.

How long the idea is expected to be effective?

While competition is emerging, there is still much room for innovation and new entrants in the market. As more companies and organizations seek to improve their security posture and protect sensitive data, the demand for passive authentication solutions is likely to continue growing. Thus, we can expect to see even more exciting developments in this field in the years to come.

BOX 12 Preconditions (optional)	BOX 13 Life cycle maintenance (optional)
Have all preconditions been met for the idea to be ready for implementation? All preconditions are met. A consortium of 4 SMEs has joined forces with the goal of creating a product that is fully prepared for the market. This partnership involves combining their resources, expertise, and technologies to create a product that meets market demands and make it ready for commercialization.	**Describe who will operate, maintain, update, and upgrade the described idea.** Consortium of 4 SMEs will take responsibility for the maintenance and support of PASID solution.

BOX 14 MISCELLANEOUS
Any additional remarks/disclaimers/comments/information you might want to provide:

3.2.2 AI-ENHANCED DISASTER EMERGENCY COMMUNICATIONS

Introduction (L3CE)

Social structures are fundamental access doors to influence societies on the short and long run. Hospitals and health care services in general is to be considered as one structure. During recent years hospitals have experienced several cyber-attacks. At the same time, the increasing number of natural disasters, experience of handling pandemic related matters, clearly indicates vulnerabilities present in health sector. In the context of hybrid threats trust in healthcare systems and proper functioning of service provision during crisis periods is essential. This threat is focused on the abilities of hospitals to provide proper services in the case of patient afflux. Such phenomena can occur in different circumstances:

- It can be natural (objective) consequences of natural (e.g.: earthquake, flood, etc.) or industrial (e.g.: leakage of hazardous substances, explosions, transport accidents, etc.) disasters as well as man maid incidents (e.g.: terrorist attack, riots, etc.).
- It can be purposefully designed (subjective) to direct extensive crowds to hospitals as a stand-alone event or part of a more sophisticated hybrid attack aiming to undermine trust in local or national healthcare system or even democracy in general.

There might be a variety of tools, methodologies and processes that can deal with different aspects of such afflux despite its nature. Those can include simulation tools, contingency planning, communication with local society, early warning, crisis management, fast mobilization of resources and many others.

Our proposed focus is made on enabling hospitals at afflux risk by any nature to make an early assessment of the situation. This can be done by having remote capabilities able to provide relevant information from the scene of action. Having initial triage, even if the afflux is purposefully designed, will allow hospitals and all parties involved in the incident handling understand what kind of incident it is and get better prepared, faster apply for resources in need and reduce direct and cascading effects.

This does not lead to the predisposition that such solution would solve all patient afflux related problems but can add capabilities in real incident handling as well as prevent afflux use in hybrid context.

BOX 1 NAME OF THE IDEA:

AI-enhanced Disaster Emergency Communications

DESCRIPTION OF THE IDEA

During major crisis the number of emergency calls has proven to be exponential, from 1 per minute to over 100 per minute, becoming impossible to sort out by emergency dispatchers, especially with the average emergency call lasting from 3 to 15 minutes dealt by just a few emergency dispatchers. Creating a massive telephonic congestion, the population is no longer capable to reach by phone the emergency services, report their positions and the evolution of their situation. This lack of communication increases the workload of Search & Rescue, which in the aftermath have to go place by place instead of focusing on population's reported positions.

The company HighWind has developed and patented the first Artificial Intelligence that can assess a patient's emergency priority level in less 100 millisecond thanks to Computer Vision and Deep learning using a crossed analysis on traumatology (nature of the wounds), emotions (pain, fears, etc.) and contextual elements (fire, smoke, etc.). Applied to major disasters, and encompassed within an smartphone "Disaster Mode" app for the population (downloaded or emulated by text-message link), it gives the emergency responders the ability to immediately visualize who are the persons most at risks on a map, prioritize search & rescue efforts to the most vulnerable persons, avoid the emergency calls congestion and facilitate patient referrals to hospitals based on the severity of their injures, thereby mitigating the potential influx of patients in hospitals.

Instead of taking one by one, lengthy emergency calls due to stressed persons, the emergency dispatch centre can perform several actions at once: send a "Disaster Mode" notification to the population, receive an accurate view on the emergency requests critical levels and positions on a map in few seconds, to better coordinate SAR efforts.

Leveraging on basic smartphone features, the AI is capable to immediately sort out victims, saving hours for the SAR teams and significantly increasing chances of survival. The "Disaster Mode" is also capable to take decisions to optimize communication based on available networks quality (no data, 2G to 5G).

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

Describe the use of the solution in reference to the gaps/need:

The solution can be used to protect the social infrastructure to potential attacks and increase resilience of health sector during the crisis situations. The solution is specifically designed to enable an early assessment of the crisis. Initial triage at the crisis scene serves the purpose of enabling hospitals and all involved stakeholders better understand the severity of crisis and prepare appropriately.

BOX 3 TYPE OF SOLUTION

Technical
Social/Human
Organizational/Process

The innovation primary pertains to technical aspects. However, to make the solution operational ready there is a need to develop a framework of utilization of the "Disaster Mode" solution and its AI-enhanced Safety Check ensuring compliance to EU General Data Protection Rules (GDPR), considering level of risks of a given disaster for the safety and health of the persons and ensure compliance of the prototype

Applicable hybrid threat domains as stated by the gaps/need: Cyber, Information, Intelligence, administration domains. **Applicable core theme(s) as stated by the gap/need:** Core Theme 1. Resilient civilians, local level, and administration.	toward EU main guidelines: AI Act, Data Act and GDPR.

BOX 4 PRACTITIONERS

Provide the applicable hybrid threat domains for which the idea is valuable: Cyber, Information, Intelligence, administration can benefit from the proposed solution.
Provide the level of practitioners in the same discipline:

- ***ministry level*** **(administration),**
- ***local level*** **(cities and regions):** Hospitals and first responders involved are the main beneficiaries of this innovation,
- ***support functions to ministry and local levels*** **(incl. Europe's third sector).**

Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments): This innovation has a potential to become highly beneficial for other first responders than health service and private citizens.

BOX 5 STATE OF THE ART

Indication of current Technology Readiness Level (TRL 1-9 index): Current TRL-6.
In which stage is the solution (research, technology, available innovation, proven innovation): The prototype of the solution is already available.
Expected time to TRL-9: 1 year.
Expected time to market: 1 year.

BOX 6 DESCRIPTION OF USE CASE(S)

This technology is very new and inventive, it was released less than one year ago, no similar technology exists in the world according to the ADIT intelligence agency, and it relies on underlying mature bricks of Computer Vision and AI.

The AI training is perform using HighWind's algorithms on medical databases, encompassing both intra and extra-hospital situations to address the three axes of detection used to qualify an emergency situation: Traumas, Context and Emotions.

The innovation can be leveraged by the companies providing emergency call reception & visualization software solutions to emergency call centres (firefighters & medical) and emergency response operators across EU.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs:

Social structures serve as fundamental channels for influencing societies in both the short and long term. Among these structures, hospitals and healthcare services, in particular, hold a significant position. Implementing an initial triage system, even when the influx of patients is moderated, enables hospitals and all relevant stakeholders to promptly identify the nature of the incident. This, in turn, facilitates better preparedness, expeditious resource allocation, and mitigation of both immediate and cascading effects.

Resilience/defensive/offensive:

The technology can be applied to enhance the resilience of social structures in the face of hybrid threats.

BOX 8 ENABLING TECHNOLOGY

Which technologies are critical in fielding the idea?

The following components are used for different modules of the prototype:

- **"Disaster Mode"** for the population's smartphones encompassing: disaster alerts reception from emergency services (type, nature, position of a disaster), pinpointing disaster on a map.
- **"Safety Check"** powered by AI for automated triage & pre-diagnostics, Emergency Calls, darken interface for battery consumption saving. Solution must be both under app and web-interface that can be spread via text-link, leveraging on existing mass-text message diffusion during disasters (e.g.: FR Alert). Safety Check feature capable upon a single-click to transmit to equipped emergency services: safety status confirmation (Safe/Unsafe), GPS position, pre-filled medical passport, pictures (selfie&main) all pre-diagnosed using HighWind's Artificial Intelligence to assess victims' critical level.
- **"Prototype an interface"** for emergency services to be able to broadcast the "Disaster Mode" on the population's smartphone, and receive a visual of the Safety Check status, triaged through AI.

BOX 9 Implementation

Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?

To make solution operational legal aspects must be considered.

- **"Disaster Mode"** solution and its AI-enhanced.
- "**Safety Check"** must be compliant to EU General Data Protection Rules (GDPR) and EU main guidelines: AI Act, Data Act and GDPR.

BOX 10 Implementation effort

Indication of costs. Describe the types of efforts and costs needed to implement the idea:

The biggest costs drivers are humans (AI engineers and lawyers) however the solution provider possess in addition the costs of AI training for the prototype, including CPU/GPU computation time which is a significant cost driver.

The total budget for the market ready solution might reach 150K EUR.

BOX 11 COUNTERMEASURES

Are there any potential countermeasures that could degrade the effectiveness of the solution?

Complex and sophisticated cyberattacks orchestrated by state actors targeting critical infrastructures and have the potential to disrupt the performance of emergency services. Nevertheless, advanced security measures are continuously evolving to mitigate such risks.

How durable is the idea (how long is the idea expected to be effective/useful)?

HighWind's AI technology is highly disruptive compared to current practices. The solution was selected, presented, and awarded during CES 2023 (Consumer Electronic Show), as one of the most disruptive technologies in service of the population. According to the ADIT intelligence firm's report and European Patent Office, no similar initiative exists in the world within research, start-up, and major companies alike.

BOX 12 Preconditions (optional)

Have all preconditions been met for the idea to be ready for implementation?

All preconditions are met. Two SMEs have joined forces with the goal of creating a product that is fully prepared for the market. This partnership involves combining their resources, expertise, and technologies to create a product that meets market demands and is ready for commercialization.

BOX 13 Life cycle maintenance (optional)

Describe who will operate, maintain, update, and upgrade the described idea.

Two SMEs: HighWind (France) https://highwind-ems.com and GAGDPR (Greece) https://gagdpr.com/EN-GB/ will take responsibility for the development maintenance and support of described idea.

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide:

End-user involvement in the solution finalization phase can add significant value.

3.3 UNDERMINE INSIDE INSTITUTIONS

EU-HYBNET responsible partner for this section: **LAUREA**

3.3.1 Advanced analytical and investigative capabilities via GRACE Platform and approach

Introduction

The threat "*Undermining institutions' internal organisation*" focuses lack of or only partly well working processes of personnel to share information in *strategic security institutions* and services and/or between different Agency/ministry/authority. Challenge also is lack of organisational agility/flexibility, mistrust between actors and culture of secrecy of strategic institutions to work together and to share the information that would support to gain more tangible results. The challenges may pave the way for hostile foreign actors to introduce, conduct and expand their malevolent actions inside the organizations that may eventually also harm society in side. It is seen that empowered cooperation between different strategic institutions is needed in order *to raise awareness about the risks inherent to organizational cultures and structures, also ways of cooperation.*

Therefore, best practices and enhanced information sharing possibilities via transparent information sharing systems/platform between *strategic institutions* or multi-Agency/ministry/authority may deliver requested solution. The use of information sharing platform provides the credit for the organization who shares the information and is expected to increase interest from others to share similar findings as well in order to have larger awareness and to solve the case(s).

BOX 1 NAME OF THE IDEA

Advanced analytical and investigative capabilities via GRACE Platform and approach

DESCRIPTION OF THE IDEA

The suggested innovation is coming from Horizon funded GRACE/ "Global Response Against Child Exploitation" project funded by the EC (GA No. 883341, duration 2020 - 2023). An important part of GRACE project has been to develop an innovative, AI-powered information sharing platform for European law enforcement authorities (LEA) investigations on child sexual exploitation and abuse material (CSEM). The AI-powered platform has answered for LEAs challenges to referrals of CSEM because the CSEM material has been in big growth on-line and the amount of material has exceeded the capacity of LEAs. The huge amount of data that needs to be analysed is a challenge in the case of hybrid threats as well and hence the GRACE projects platform and way how to share sensitive data is seen as a promising solution for authorities who are dealing analysis of signs of hybrid threats

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

Applicable hybrid threat domains as stated by the gaps/need:

Applicable core theme(s) as stated by the gap/need:

The novel way how to share the data in the GRACE platform is such that it may decrease the barriers of strategic institutions to share information between each others, also among the personnel in their own institutions. The GRACE platform also provides clear procedure for cooperation and credit for work done. This may decrease the culture of secrecy and hence leave less room for adversaries' influence inside the strategic institution and/or between institutions.

BOX 3 TYPE OF SOLUTION

- **Technical**

As defined in GRACE platform, the platform should be tailored to following services.

1) Deliver a semi-automated content analysis and prioritisation mechanism that enables authorities to better apprehend offenders
2) Implement a secure and privacy-aware, federated learning infrastructure for beyond state-of-the-art content classification across Europe.
3) Facilitate authorities ability to track short and long-term trends in the production of hybrid threats campaigns. This is to alert for preventive strategies and policy decisions

- **Social/Human**
- **Organizational/Process**

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:** The main beneficiaries would be strategic institutions analysing hybrid threats in all the domains where hybrid threats may occur.

- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
    The tool is especially relevant for strategic institutions analysis.
  - **II)** ***local level*** **(cities and regions):**
    If there are strategic institutions in local level, the solution serves also them.
  - **III)** ***support functions to ministry and local levels*** **(incl. Europe's third sector):**
    If there are strategic institutions in third sector, the solution serves also them.

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
- The end users would e.g. intelligence to analyse large amount of data and to connect dots together in order to discovered hybrid campaign. Also, to find similar cases in order to find possible patterns of malicious actors.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  TLR 7-8
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  The technology to be used for innovation already exists in GRACE. There would be need to tailor the solution for strategic institutions use to analyse hybrid threats related material.
- **Expected time to TRL-9.**
  Up to 5 years.
- **Expected time to market.**
  Up to 5 years.

BOX 6 DESCRIPTION OF USE CASE(S)

Solution is seen beneficial in following scenario.

Scenario. In country X LEAs in city X1 and city X2 and city X3 report that amount of youngster gangs has increased suddenly and they target to harm police. The same issue is reported by Country Y in city Y1 and city Y2; also in country Z, LEAS in city Z1-Z5 report about the increase of same phenomenon. At the same time it has been recognized that in country X and Y cyber-attacks and espionage has been made to local police stations in order to solve where and how the police are patrolling in certain areas of city X1-X3 and city Y1-Y2. Furthermore, from Country X and Country Z it has been reported that there are disinformation campaigns on police's violence against youth gangs and the campaigns promotes youth gangs to attack to police. In addition, from Country Y and Country Z it is reported that foreign direct investments have been increasing to Youngsters Homes where agitation of youngster to youngster gangs often take place too. The same FDI has been then noticed also in country Z. To connecting all these dots together, countries Z, Y, Z consider that they are under similar influence campaigns where goal is to agitate youngster to join youngster gangs and to harm the society's stability by attacking police that would further provide image of instability in the society and lead to the discussion is the society taking care of youngsters and should some new political decision to be made in order to get the situation in control.

Due to analysis and reporting done on the named phenomenon in the suggest information sharing platform, similar increase of youngster gangs and FDI to Youngster Homes are discovered in country A and country B. The LEAs in City X1-3 alike in City 1-2 are credited on sharing their notion in their country but also to other countries because otherwise the other actions related to the same phenomenon might not have been revealed.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**
  The solution provides new AI-powered platform to go through signs of hybrid threats and to deliver analysis of the cases and signs. This support to empower cooperation between strategic institutions information sharing when the goal to find the malicious means and operations are seen as a joint action. The discoveries are also credited to the organization who shares them and the traceability provides also transparency leaving less room for adversaries.
- **Resilience/defensive/offensive**
  The innovation contributes to offensive measures. However, when the solutions supports to discover patterns of malicious means, the solution will also in the long run also to support increase of resilience in society and defence to similar cases.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  Artificial Intelligence (AI).

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
- Security. If malicious actors may have access to the platform they may learn about the analysis.
- New solution may feel burdensome for uptake next to existing systems. Because the solution would be used by many strategic institutions, legal aspects for information sharing should be solved.

<table>
<tr>
<td>BOX 10 Implementation effort
- Indication of costs:
Describe the types of efforts and costs needed to implement the idea.
The funding of GRACE project has been c. 7.000.000 Euro and the same is expected to the development project focusing on delivering similar platform for authorities focusing on hybrid threats.</td>
<td>BOX 11 COUNTERMEASURES
- Are there any potential countermeasures that could degrade the effectiveness of the solution?
The malicious actors may use cyber-attacks to harm the solution and to learn on its content.
- How durable is the idea (how long is the idea expected to be effective/useful?)
- The solution is durable as long as the AI in the platform is updated regularly.</td>
</tr>
<tr>
<td>BOX 12 Preconditions (optional)
- Have all preconditions been met for the idea to be ready for implementation?
- Discussion between strategic institutions interest for such a tool should be initiated. To support this, opinions from LEAs on the use of GRACE platform should be heard.</td>
<td>BOX 13 Life cycle maintenance (optional)
- Describe who will operate, maintain, update, and upgrade the described idea.
The solution provider or if the solution is handed over to strategic institution to run, then it is in their responsibility.</td>
</tr>
<tr>
<td colspan="2">BOX 14 MISCELLANEOUS
Any additional remarks/disclaimers/comments/information you might want to provide

Uptake of new platform might feel burdensome and hence the benefit to use the solution should be well explained for the strategic institutions. The platform should justify that the analysis it provides cannot be gained otherwise and the analysis is to trust on. This set high demands to the AI in the platform. Strategic institutions should be deeply involved with the development process in order to gain solution that the strategic institutions wish to use in their cooperation. The solution could be used not only by national strategic institutions but also EU level strategic institutions who are sharing information on hybrid threats campaigns and means.</td>
</tr>
</table>

3.3.2 'Antidote' to hostile messaging delivered by private messaging apps

Introduction

The integrity of organizations can mainly through the people involved with the organization – mainly of course the staff members, but also the 'clients' (e.g. students in the case of educational establishment, patients in the case of hospital, local inhabitants in the case of municipality, criminals in the case of police, etc.), stakeholders and others. Sawing distrust between these people is one of the most efficient hostile modus operandi to undermine institutions' internal organization. In most cases the circle of people involved is much wider than just the staff members of the organization, sometimes even (local) community as a whole.

In order to approach the people involved with organizations different avenues may be used by hostile actors and it is difficult to close – in many cases illegal or counterproductive – such avenues. Another option would be immunizing the people in question for hostile messages. This innovation aims to use one of the hostile avenues of approach – the private messaging applications – for such immunization. Although the 'antidote' is 'injected' through one channel, if it works, it should also block the other channels used by hostile entities.

BOX 1 NAME OF THE IDEA

'Antidote' to hostile messaging delivered by private messaging apps

DESCRIPTION OF THE IDEA

Integrity of organizations can be undermined through the people involved with the organization – the circle much wider than only the staff members of the organization. Currently, one of the main channels where hostile messages, disinformation, conspiracy theories, etc. reach people are private messaging apps. This innovation proposes using the same channel for preventive work and reaction. Although currently Telegram has been most frequent private messaging app to be used for spreading disinformation, the innovation should keep in mind also the future and other apps (Signal, Viber, WhatsApp, etc.).

The information to be shared in order to immunize would need to raise the awareness and standard of critical thinking. I.e. messages like "this is not true" may not be the most efficient, but rather the games attracting attention to the problem may be used. The technically most simple solutions are sharing the link to freely available and already existing games.

'Antidote' can be shared in two major ways. The simplest, but also much more expensive way, is to buy it as advertisement. The more complex, but much cheaper and more efficient way, would be to collaborate with the owners of the private messaging apps in order to sort out the target groups to be immunized and share the content for free.

Although the best 'antidote' could be chosen by the organization which integrity needs protection, the communication and dealing with private messaging app owners should be handled centrally.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  Improving the resilience to hostile messaging of people and therefore fostering the integrity of organizations.

- **Applicable hybrid threat domains as stated by the gaps/need:**
  Political, Social, Informational

- **Applicable core theme(s) as stated by the gap/need:**

BOX 3 TYPE OF SOLUTION

- **Technical**

- **Social/Human**

- **Organizational/Process**
  The technical solution – private messaging apps – is already there. There is only a need to start using them more efficiently in the fight countering disinformation and getting the owners of the app on board.

BOX 4 PRACTITIONERS

Provide the applicable hybrid threat domains for which the idea is valuable: Political, Social, Informational

- **Provide the level of practitioners in the same discipline:**
  - **I) *ministry level* (administration):**
  - **II) *local level* (cities and regions):** municipalities, education establishments, social welfare organizations, law enforcement, etc.
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):** NGOs, private companies.

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  The main beneficiaries would be all organizations of democratic societies in direct interaction with local inhabitants: municipalities, education establishments, social welfare organizations, NGOs, law enforcement and others.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9

- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  9

- **Expected time to TRL-9.**
  0y

- **Expected time to market.**
  Technological solution exists – the innovation is about how to better use it.

BOX 6 DESCRIPTION OF USE CASE(S)

The internal integrity of the organization is under attack by hostile messaging, disinformation, conspiracy theories, etc. Not only the employees of the organization but also outside stake holders are targets of hostile messaging and they put additional pressure to the organization and create serious problems for the organization that may cause its integral structure disintegrating.
In such conditions the organization starts to spread 'antidote' via private messaging apps. The organizations in question may be (not limited to):

- Municipalities
- Education establishments
- Law enforcement
- NGOs
- Private companies

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

- **Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.**

- **Resilience/defensive/offensive**
  Resilience, defensive

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  The private messaging apps; the games developed for rising awareness and critical thinking.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  The owners of the private media apps may be reluctant to accept the idea, referring to the trust of the consumers. Of course all data protection regulations have to be followed.

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.
  Depends on the possible deal with the owners of the private messaging app owners. Can be for free (i.e. only the work time of negotiators spent) or range to multi-million bills if the 'antidote' has to be distributed as advertisement.

BOX 11 COUNTERMEASURES

Are there any potential countermeasures that could degrade the effectiveness of the solution?
Hostile propaganda and cyberattacks, theoretically also electronic jamming of smartphones or wireless internet.

- **How durable is the idea (how long is the idea expected to be effective/useful?)**
- Depends on the future of private messaging applications and the content of the deal with their owners.

BOX 12 Preconditions (optional)	BOX 13 Life cycle maintenance (optional)
- **Have all preconditions been met for the idea to be ready for implementation?**	- **Describe who will operate, maintain, update, and upgrade the described idea.** Depends on the possible deal with the owners of the private messaging app owners. Can be for free (i.e. only the work time of negotiators spent) or range to multi-million bills if the 'antidote' has to be distributed as advertisement.
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide**	

4. INNOVATIONS FOR COUNTERING HYBRID THREATS: CORE THEME: INFORMATION AND STRATEGIC COMMUNICATIONS

4.1 MEDIA CONUNDRUM

EU-HYBNET responsible partner for this section: **KEMEA (Threat 15,8)- ZITiS (11,12)**

4.1.1 MEDIA PLURALISM MONITOR (MPM)

Introduction

Journalistic media is faced with an environment of increased competitiveness today, which is not to its advantage. Countering disinformation, fact checking, as well as straightening up the debate and public discourse, all induce high costs. The added factor of competition amongst media outlets who rely on click-bait practices consequently leads to depriving market shares from quality journalistic media. The identified gap therefore is the lack of true quality journalistic competitiveness (i.e. journalistic integrity), as opposed to competition in terms of which outlet will first transmit the story (i.e. speed of news coverage), sacrificing in this way any sufficient investment in investigative journalism. The need which arises therefore is the identification and sharing of best practices for journalistic media economic sustainability. A correct and accurate situational snapshot of the ways in which media can be sustained is required, to strengthen the economic model of journalism. Creating a register of good and best practices whereby media outlets have increased their economic viability without compromising content quality and journalistic investigations would therefore be required, across the EU.

Innovation Recommendation

The Media Pluralism Monitor (MPM) is a tool developed by the Centre for Media Pluralism and Media Freedom (CMPF) of the European University Institute (EUI) to assess the potential weaknesses in national media systems that may hinder media pluralism.

Based on 20 indicators, summarizing 200 variables, it covers four areas:

1. Fundamental protection
2. Market plurality
3. Political independence
4. Social inclusiveness.

The news media industry has been severely hit by the COVID-19 pandemic and the accompanying economic crisis. Although the shock was largely foreseeable due to the extraordinary circumstances, its depth and the diverging effect between different countries has to be investigated.

The MPMCOVID-19 data collection offers some provisional results, that can be summarised as follows:

- In the 18 EU countries covered by the research, revenues for news media have experienced an average decrease that was beyond the GDP decrease during the first wave of the COVID19 pandemic.

- The decline is driven by the fall in the advertising expenditure, which decreased by more than the percentage of GDP decrease, in all the European countries covered by the research. The shift to digital, which was experienced in consumer behaviour (including an increased willingness to pay for online content), and reflected in business strategies, has not managed to counterbalance the decrease in advertising income.

- Different trends were visible in different news media sectors. Legacy television and radio stations, whose business models substantially rely on advertising, were severely hit, whereas video-on-demand and video-platforms benefited from the surge in subscriptions. The long-lasting crisis of newspapers and local media has worsened, as they suffered hits from both sides, advertising and print sales, with the digital subscription not making up for the losses. Closures and reductions in the news media outlets have not been massive in the first wave of the pandemic, partly due to extraordinary measures (see below), but are foreseen in several countries, particularly for newspapers and local media. The digital news media performed relatively better, particularly the ones based on pay-models rather than on advertising.

- A number of news outlets started experimenting with alternative revenues - especially some form of reader-generated revenue online (such as donations, subscriptions or membership). While there is no comprehensive data that would allow us to estimate the role these revenues play in the respective media landscapes, the experiences shared with us about some successful digital media efforts give some grounds for optimism. However, reader revenues (especially subscriptions to paywalled material) have their own risks. In the short term, locking one's content can hurt the page's SEO ranking, readership and advertising revenues. Not to mention, that it raises a serious ethical dilemma: if less content will be available for free on the internet, low-income people as well as those who are less willing to pay for content will be cut off from quality content and more at risk of falling prey to misleading content.

- Employment and salary trends highlight a high risk for the economic safety of journalists, despite the surge in demand for information and trusted sources. In this case, differences may be traced with regard to the contractual status of journalists, rather than to the sector in which they work. Journalists employed in the newsrooms on a regular basis have been relatively protected by the widespread use of job retention schemes, suffering eventually from salary cuts, but still keeping their jobs. But the support programs often fell short of covering all the journalists, protecting at a smaller extent, or not protecting at all, the freelancers and the journalists with non-standard contracts. The economic situation of freelancers worsened all over the countries covered by the research.

- Public support was available in the overwhelming majority of the countries covered, even including extraordinary subsidies, but the amounts provided to the media were often seen as insufficient in light of the much larger pandemic-induced losses.

In this context the work carried out suggests a series of possibilities to improve the viability of the EU media landscape in the post COVID-19 era.

- The news media sector should be included in the national recovery and resilience plans, with an emphasis on incentivising investments that support the transition towards digital news media.

- Labour and social policies should eliminate or mitigate the dualism in the journalistic labour market. Considering also the generational gap (employed journalists are often older, freelancing and non-standard contracts are more common in younger generations), there is a huge risk of precariousness and vulnerability for the next generation of journalists, which in turn may menace the safety of journalists (Zuffova and Carlini, 2021) and threaten freedom of information. A universal protection scheme for all journalists, regardless of their contractual status, should be studied and implemented.

- Public subsidies will remain relevant, thus policy makers need to work on subsidies that are sufficient and follow the most appropriate approach. These subsidy schemes need to be based on transparent criteria to avoid capture, able to react to the changing conditions in the media environment, avoid distorting the market or creating dependence, and put an emphasis on journalism that contributes to the public good.

- Innovation should be a priority, both in news media production as well as in business models, in order to avoid the risk of financing outdated models that were already in deep crisis before COVID-19.

BOX 1 NAME OF THE IDEA

Media Pluralism Monitor (MPM)

DESCRIPTION OF THE IDEA

Media Pluralism Monitor (MPM) is a tool developed by the Centre for Media Pluralism and Media Freedom (CMPF) of the European University Institute (EUI) to assess the potential weaknesses in national media systems that may hinder media pluralism.

Based on 20 indicators, summarizing 200 variables, it covers four areas:

1. Fundamental protection
2. Market plurality
3. Political independence
4. Social inclusiveness.

The news media industry has been severely hit by the COVID-19 pandemic and the accompanying economic crisis. Although the shock was largely foreseeable due to the extraordinary circumstances, its depth and the diverging effect between different countries has to be investigated.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  The solution can be used to prevent the deprivation of market shares from quality journalistic media by ensuring that sufficient investment in investigative journalism is not sacrificed in the face of journalistic competitiveness, by identifying and sharing best practices for journalistic media economic sustainability.

- **Applicable hybrid threat domains as stated by the gaps/need:**
  Information, cyber, culture, social/societal, legal, economy, intelligence domains could benefit from such solutions.

- **Applicable core theme(s) as stated by the gap/need:**
  Information and Strategic Communications

BOX 3 TYPE OF SOLUTION

- **Technical**

- **Social/Human**
  The innovation proposed is a Social/Human one

- **Organizational/Process**
  Whereas the innovation proposed is a Social/Human one, it entails compliance to various policies, therefore making the necessity of legal requirements and a legal framework substantial.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Information, cyber, culture, social/societal, legal, economy, intelligence domains could benefit from such solutions.

- **Provide the level of practitioners in the same discipline:**
  The threat was listed under Civic *Space* in Deliverable D2.7 (Long List of Gaps and Needs)
  - I) ***ministry level*** **(administration):**
  - **II) *local level* (cities and regions):**

- **III) *support functions to ministry and local levels* (incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Media outlets, journalists, publishers, broadcasters, editors and other related stakeholders are the end-users of the idea.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9

- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  Available innovation

- **Expected time to TRL-9:**
  0 years

- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

Media Pluralism Monitor (MPM) assesses the potential weaknesses in national media systems that may hinder media pluralism and covers the areas of fundamental protection, market plurality, political independence and social inclusiveness.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.

As mentioned above, the solution can be used to help mitigate the loss of market shares from quality journalistic media (i.e. sufficient investment in investigative journalism), by bridging the gap between true quality journalistic competitiveness (i.e. journalistic integrity), as opposed to competition in terms of which outlet will first transmit the story (i.e. speed of news coverage). The need which arises therefore is the identification and sharing of best practices for journalistic media economic sustainability. A correct and accurate situational snapshot of the ways in which media can be sustained is required, to strengthen the economic model of journalism. Creating a register of good and best practices whereby media outlets have increased their economic viability without compromising content quality and journalistic investigations would therefore be required, across the EU. The solution proposed falls under the information and strategic communications theme.

- **Resilience/defensive/offensive**
  The solution can be used in countering hybrid threats by promoting journalistic economic viability by ensuring its resilience, by defending against threats towards journalistic economic viability.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  N/A

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No.

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  No.

	- **How durable is the idea (how long is the idea expected to be effective/useful?)** Since media in all its forms will exist indefinitely, the solution is expected to be useful for a very long time.
BOX 12 Preconditions (optional) - **Have all preconditions been met for the idea to be ready for implementation?** No preconditions exist.	**BOX 13 Life cycle maintenance (optional)** - **Describe who will operate, maintain, update, and upgrade the described idea.**
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide**	

4.2 SECTARIANISM

EU-HYBNET responsible partner for this section: KEMEA

4.2.1 "BAD NEWS" PREBUNKING GAME PLATFORM

Introduction

Information operations - regardless of whether they are labelled as political warfare or influence operations or exercised as an element of a broader hybrid campaign - exploit the vulnerabilities of modern democracies and target both the elites and societies of the western states and specifically identity minorities in order to influence political behaviour and public opinion. The hostile toolkit involves the dissemination of false, misleading, and manipulative information in the media - especially the social media. Information operations exploit one of the most challenging characteristics of our era: ambiguity. The lines between virtual and real, domestic, and international, public and private, normality and disorders, have eroded, and the result is far more ambiguity. Planting and disseminating a lie via social media is cheap and easy. On the other hand, identifying the lie, tracking its origins, and communicating 'your' truth to the same audiences is labour intensive and costly.

Public debate is thus characterized by a multitude of identities, we-feelings, and In-groups built on the assumption they would be victims of persecution or discriminatory treatment and hate speech. This style of communication creates new norms of behaviour, language, acceptability regimes threatening the existing social cohesion in some cases. The use of this communication style can sow instability through foreign interference campaigns.

Hate speech is prevailing, especially deriving from these victimized minorities that further provoke responses (sometimes violent) from the population majority. It is observed that there is a lack of productive communication and dialogues in the public sphere.

It is important to assess discourses that can carry potential for violence and to link those with a registry of harmful content. Whoever considers and proclaims himself as a victim, usually is producing hate speech and this recycles the phenomenon from the other side as well, promoting social cleavages and instability.

Concern about people's general vulnerability to political indoctrination goes back many decades, arising at the time from disquietude about persuasive techniques employed by totalitarian states. The larger question of how to go about developing attitudinal "resistance" against unwanted persuasion attempts ultimately led McGuire[25] to develop "inoculation theory", which, for a popular audience, he described as a "vaccine for brainwash" which led nowadays to the "Prebunking Approach".

[25] McGuire, W. J., Resistance to Persuasion conferred by Active and Passive prior Refutation of the Same and Alternative Counterarguments, The Journal of Abnormal and Social Psychology, 63(2), 326–332, 1961.

BOX 1 NAME OF THE IDEA

"BAD NEWS" Prebunking Game

DESCRIPTION OF THE IDEA

Active prebunking interventions require the individual to take action, making choices that help them retain information and engage more deeply with the content they see. The primary active approach researched to date is games[26]. While games are more immersive and allow individuals to be inoculated against multiple manipulation techniques commonly used in misinformation, they require a larger investment from the viewer in terms of time and focus, which may reduce the number of people engaging with it. They are also a larger investment to produce, though some high-impact games have been implemented on a large-scale. The inoculation metaphor relies on a medical analogy: by pre-emptively exposing people to weakened doses of misinformation cognitive immunity can be conferred. An example is the Bad News game, an online fake news game in which players learn about six common misinformation techniques. (https://www.getbadnews.com/books/english/ offered in 23 languages)

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  This was the first-ever prebunking game. It is a choice-based browser game created by DROG and the University of Cambridge in which players take on the role of a fake news producer and learn to identify and mimic six misinformation techniques (e.g. trolling, conspiratorial reasoning, impersonation) over six levels. Since then, several other games with similar premises have been designed.
- **Applicable hybrid threat domains as stated by the gaps/need:**
  Cyber, Information and Military /Defense domains could benefit from such solutions.
- **Applicable core theme(s) as stated by the gap/need:**
  Cyber and Future Technologies

BOX 3 TYPE OF SOLUTION

- **Technical**
  The innovation proposed is a technical one.
- **Social/Human**
  Broadly speaking, the longer and more engaged the person is with the prebunk, the greater the size and duration of the prebunking effect.
- **Organizational/Process**

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Cyber, Information and Military/Defence domains could benefit from such solutions.
- **Provide the level of practitioners in the same discipline:**
  The threat was listed under *Services Space* in Deliverable D2.7 (Long List of Gaps and Needs)
  - I) ***ministry level*** **(administration):**
  - **II) *local level* (cities and regions):**
  - **III) *support functions to ministry and local levels* (incl. Europe's third sector):**
- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Citizens, users, news consumers, audiences in general

[26] Basol, M., Roozenbeek, J., and Van der Linden, S., "Good News About Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News," Journal of Cognition 3, no. 1, 2020.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  9

- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  Available innovation

- **Expected time to TRL-9:**
  0 years

- **Expected time to market.**
  0 years

BOX 6 DESCRIPTION OF USE CASE(S)

It can fragment videos, to enhance and explore key frames and images through a magnifying lens, to read video and image metadata, to check media copyright licenses, and to apply forensic filters on still images, all of which are vital tools for anyone trying to check the authenticity and source of the material.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.

Playing Bad News significantly improves people's ability to spot misinformation techniques compared to a gamified control group, and crucially, also increases people's level of confidence in their own judgments. Importantly, this confidence boost only occurred for those who updated their reliability assessments in the correct direction. This offers further evidence for the effectiveness of psychological inoculation against not only specific instances of fake news, but the very strategies used in its production. Implications are supported for inoculation theory and cognitive science research on fake news.

- **Resilience/defensive/offensive**
  The solution can be used in countering hybrid threats in all three manners: to promote resilience, to defend against a threat and also to provide offense against a threat.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  The purpose of the game is to produce and disseminate disinformation in a controlled environment whilst gaining an online following and maintaining credibility. Players start out as an anonymous netizen and eventually rise to manage their own fake news empire. The theoretical motivation for the inclusion of these six strategies are explained in detail in Roozenbeek and van der Linden (2019) [27] and cover many common disinformation scenarios including false amplification and echo chambers. Moreover, although the game scenarios themselves are fictional they are modelled after real-world events. In short, the gamified inoculation treatment incorporates an active and

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No.

[27] Roozenbeek, J., & van der Linden, S., Fake News Game confers Psychological Resistance against online Misinformation, Nature Palgrave Communications, 5(65), 2019.

<table>
<tr><td>experiential component to resistance-building.</td><td></td></tr>
<tr><td>BOX 10 Implementation effort
- Indication of costs:
Describe the types of efforts and costs needed to implement the idea.
The solution is free of charge.</td><td>BOX 11 COUNTERMEASURES
- Are there any potential countermeasures that could degrade the effectiveness of the solution?
Progress in disinformation methodologies.

- How durable is the idea (how long is the idea expected to be effective/useful?)
The solution is expected to be useful for a very long time.</td></tr>
<tr><td>BOX 12 Preconditions (optional)
- Have all preconditions been met for the idea to be ready for implementation?
No preconditions exist.</td><td>BOX 13 Life cycle maintenance (optional)
- Describe who will operate, maintain, update, and upgrade the described idea.
Exclusive licensing of this game for purposes of scientific research has been granted to researchers at the Cambridge Social Decision-Making Lab at Cambridge University. This website and its contents may only be used with the permission of TILT and the Cambridge Social Decision-Making Lab. Tilt is a company that aims at increasing people’s information resilience and arming them against the harmful influences of disinformation and online manipulation. This game was designed and developed by Gusmanson, a design studio creating serious games.</td></tr>
<tr><td colspan="2">BOX 14 MISCELLANEOUS
Any additional remarks/disclaimers/comments/information you might want to provide
It will be important for future research to evaluate to what extent “active” gamified inoculation is superior to “passive” approaches—including traditional fact-checking and other critical thinking interventions—especially in terms of eliciting a) motivation, b) the ability to help people discern reliable from fake news, and c) the rate at which the inoculation effect decays over time.</td></tr>
</table>

4.3 ATTACK ON INFORMATION

EU-HYBNET responsible partner for this section: **ZITiS**

4.3.1 REAL-TIME FACT-CHECKING BROWSER EXTENSION

Introduction:

The use of independent fact-checkers aims to promote accuracy and accountability in the dissemination of information. Different results are expected when using it.
First, increased accuracy, with independent fact-checkers responsible for verifying the factual accuracy of claims made by individuals, organizations, or media outlets. By providing reliable and evidence-based assessments, they help ensure that the information provided to the public is accurate and reliable.
In addition, credibility is enhanced as fact-checkers play a crucial role in enhancing the credibility of news sources and information providers. When independent fact-checkers confirm a particular claim or refute misinformation, it increases the credibility of the source and helps individuals make more informed decisions.
Misinformation and disinformation are quickly caught and contained. Misinformation and disinformation often spread quickly via social media and other online platforms. Independent fact-checkers act as a defence against the spread of misinformation by identifying and uncovering inaccuracies and untruths. This helps prevent the negative impact of misinformation on democratic processes, public opinion and health.
Individuals as well as organizations receive increased transparency. Fact-checkers typically use transparent methods, citing sources and providing evidence to support their findings. This transparency is the basis for understanding all of the fact-checkers' conclusions and making one's own judgments based on the evidence presented.
Constantly increasing and sharpening public awareness and education through independent fact-checkers helps to enable critical thinking, media literacy and fact-based decision-making. You help them educate the public about the value of information sources by uncovering the methods of spreading misinformation and becoming more sophisticated consumers of news and other content.

Overall, the expected result of using independent fact-checkers is to enable a better informed and critically engaged public, reduce the impact of misinformation, and provide a stronger basis for evidence-based discussions and decisions.

BOX 1 Name of the solution

Real-Time Fact-Checking Browser Extension

Description of the solution

A constant companion when surfing the Internet can be a revolutionary browser extension, perhaps even with the most modern AI algorithms. An innovation in this direction can lead in a new era in the fight against misinformation and false claims.

As you surf the vastness of the Internet, this browser extension carefully scans the content of the web pages you visit, leaving no text unedited and saving images or videos for careful viewing.

Basically, it uses advanced Natural Language Processing (NLP) algorithms to ensure that every claim, statistic, or factual statement that is detected goes through a rigorous fact-checking process. If there are doubts about a claim, the browser plug-in will react quickly.

The technology is not just limited to text, but extends its perception to images and videos, using advanced image and video analysis to detect potentially manipulated or misleading media content. It is a guard that protects against visual illusion.

The secret to its incredible accuracy lies in its extensive database of verified information, fact-checking reports and reliable sources. It compares the claims against this deposit and provides you with the most accurate and up-to-date information available. When the plugin detects a fact-checked claim that turns out to be false or misleading by scanning the digital landscape in real time, it will gently but firmly alert you with a pop-up. This alert not only highlights the correct information, but also provides a direct link to a reliable fact-checking source.

To make this plugin even more powerful and responsive to user needs, it includes an extensive verification component. Users have the right to report claims they find questionable, encouraging further investigation and community collaboration in the search for the truth. What about privacy? The extension puts everyone's privacy first and treats other browsing data with the utmost care. It captures and stores only what is relevant to the fact-checking process, while respecting the digital footprint.

Essentially, this real-time AI fact-checking browser extension gives the internet user a powerful tool against misinformation. It encourages critical thinking, keeps the user well-informed and, most importantly, ensures that truth prevails in the digital world.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**
  The solution can be used to improve the trust in correctness of news and helps against spreading of fake news.
- **Applicable hybrid threat domains as stated by the gaps/need:**
  Civic space, governance space, services space could benefit from such solutions.
- **Applicable core theme(s) as stated by the gap/need:**
  Information and strategic communications

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

The innovation is of technical nature combined with Social/Human theme. A legal framework would also be necessary in countering the problem and its routes as well as the civilians needs to be anxious to use these methods.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Civic space, governance space, services space could benefit from such solutions.
- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
  - **II) *local level* (cities and regions):**
  - **III) support functions to ministry and local levels (incl. Europe's third sector):**

- **Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):**
  Private citizens and governance space are the main beneficiaries of these kinds of technologies.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  6
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  The most important tasks have to be done in the parts research ,technology and bring to market.
- **Expected time to TRL-9.**
  2-4 years, an end of improvements will never be reached.
- **Expected time to market.**
  2-4 years with an important work for advertising for using the functionality.

BOX 6 DESCRIPTION OF USE CASE(S)

This technology and tools can be used in every situation of the whole everyday life in every situations of all men who consuming information and news in all formats, e.g. Text, audio or video.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.
As mentioned above, these technologies can help all individuals and it can help societal resilience.

- **Resilience/defensive/offensive**
  The idea can be used in a defensive and offensive way to counter hybrid threats.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  The databases which are used for the proven facts needs to be verified all the time and needs to be certified.
  Faked browser plugins could be produced and offered which doing false fact checks.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No.

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.
  1-2 Million €/a until the base system is ready after some regular costs for maintenance and improvement.

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  Wrong verified facts stored in the database lead to false positive checked news.
- **How durable is the idea (how long is the idea expected to be effective/useful?)**
  The solution is expected to be useful for a very long time, with a steady maintenance and efforts for monitoring the newly added or changed entries to many services of social media, news platforms and maybe messengers.

BOX 12 Preconditions (optional)

- **Have all preconditions been met for the idea to be ready for implementation?**
  No preconditions are needed.
  Current browsers are supporting all necessary techniques for this.

BOX 13 Life cycle maintenance (optional)

- **Describe who will operate, maintain, update, and upgrade the described idea.**
  The services shall be offered by more than one operator. Maybe founded installations and/or commercial instances from e.g. news agencies who are interested in the correctness of their reports, pictures/videos or audio streams.

BOX 14 MISCELLANEOUS

Any additional remarks/disclaimers/comments/information you might want to provide

4.3.2 BLOCKCHAIN -BASED VERIFICATION

Introduction

The increased use of visual illusions can have a number of negative consequences. Fake videos, deep fakes, and manipulative images that are quick to spread false information can mislead the public and skew their opinions. People's trust in the media, institutions and even each other is decreasing because they are constantly exposed to visual misinformation. Because of this loss of trust, people find it difficult to differentiate between authentic and manipulated media, contributing to an overall loss of trust.

Visual misinformation can be used as a tool to polarize society, exacerbate existing divisions, and manipulate public opinion. It is easier to manipulate people's beliefs and influence their perceptions by deliberately disseminating misleading images that support certain narratives or ideologies. Using visual misinformation can also seriously damage the reputations of people, businesses, and even entire communities. Altered images and videos can be used to falsely accuse people of damaging a company's reputation or creating public scandals. Even if the misinformation is later debunked, the damage to reputation can be lasting.

The impact of increasing levels of visual misinformation not only affects a person's reputation, but also changes the social and political climate. It undermines democratic processes, distorts elections and fuels social unrest. False or manipulated images can incite violence, trigger outrage and provoke conflict by exploiting people's emotions.

The spread of visual misinformation also poses challenges for media companies and technology platforms responsible for moderating content. In order to detect and combat the proliferation of misleading images, efficient mechanisms and algorithms must be created, which can be difficult and time-consuming.

Prioritizing media literacy, critical thinking, fact checking, and responsible use of visual content is critical to preventing these potential consequences. By developing these skills, people can navigate the world of visual information more successfully. Fully addressing the challenges associated with visual misinformation will also require collaboration between technology companies, policymakers and society at large.

BOX 1 NAME OF THE IDEA

BLOCKCHAIN-BASED VERIFICATION

DESCRIPTION OF THE IDEA

Blockchain technology can play a crucial role in the fight against the increasing use of visual misinformation. By leveraging the inherent security and transparency of blockchain, a robust system can be established to verify the authenticity of images and videos. Blockchain allows us to timestamp visual content at the time of creation. Each medium is linked to a unique cryptographic hash and recorded on the blockchain, creating an immutable record of its provenance. This timestamp ensures that the authenticity of the content can be easily verified, thus helping to identify real footage and distinguish it from manipulated images.

In addition, the blockchain can store tamper-proof metadata about the content, such as information about the author, location and any editing history, further strengthening the credibility of the information from the visual media.

Decentralized content sharing platforms are built on blockchain technology, which ensures that content is verified before being widely distributed. Mobile apps and browser plug-ins to be developed, will enable users to use blockchain-based verification. Such tools allow people to easily examine the blockchain records of visual content they reach online.

Fact-checking organizations are integrating this blockchain technology into their processes by recording their findings and conclusions on the blockchain. This creates an immutable record of verified information, increasing confidence in their reviews. Collaborating with content creators is essential. Encouraging professionals and journalists represents a sign of trustworthiness to certify the authenticity of their work on the blockchain. This also increases trustworthiness in a time plagued by misinformation. Public blockchain visual content verification databases managed by a consortium of organizations can further improve transparency and accountability. Furthermore recognizing blockchain as evidence in court cases related to misinformation is an incentive to use this technology to verify content.

It is important to note that while blockchain can improve the trustworthiness of visual content, it is not a panacea. It has its limitations and its effectiveness depends on widespread acceptance and proper implementation. Additionally, blockchain should be used in conjunction with other strategies such as media literacy and fact-checking to create a comprehensive approach to combating visual misinformation.

BOX 2 REFERENCE TO CAPABILITY GAPs/NEED

- **Describe the use of the solution in reference to the gaps/need**

  The solution can be used to improve the trust in correctness of news and helps against spreading of fake news.

- **Applicable hybrid threat domains as stated by the gaps/need:**

  Civic space, governance space, services space could benefit from such solutions.

Applicable core theme(s) as stated by the gap/need:

Information and strategic communications.

BOX 3 TYPE OF SOLUTION

- **Technical**
- **Social/Human**
- **Organizational/Process**

  The innovation is of technical nature combined with Social/Human theme. A legal framework would also be necessary in countering the problem and its routes as well as the civilians needs to be anxious to use these methods.

BOX 4 PRACTITIONERS

- **Provide the applicable hybrid threat domains for which the idea is valuable:**
  Civic space, governance space, services space could benefit from such solutions.

- **Provide the level of practitioners in the same discipline:**
  - I) ***ministry level*** **(administration):**
  - **II)** ***local level*** **(cities and regions):**
  - **III) <u>support functions to ministry and local levels (incl. Europe's third sector):</u>**

Provide the end-users of the idea (such as NGO's, private citizens, private companies, media outlets, police, firefighting departments):
Private citizens and governance space are the main beneficiaries of these kinds of technologies.

BOX 5 STATE OF THE ART

- **Indication of current Technology Readiness Level (TRL 1-9 index)**:
  6
- **In which stage is the solution (research, technology, available innovation, proven innovation):**
  The most important tasks have to be done in the parts research ,technology and bring to market.
- **Expected time to TRL-9.**
  2-4 years, an end of improvements will never be reached.
- **Expected time to market.**
  2-4 years with an important work for advertising for using the functionality.

BOX 6 DESCRIPTION OF USE CASE(S)

This technology and tools can be used in every situation of the whole everyday life in every situations of all men who consuming information and news in all formats, e.g. text, audio or video.

BOX 7 IMPACT ON COUNTERING HYBRID THREATS

Describe how the idea contributes to countering hybrid threats; relate this to one or more capability gaps and needs.
As mentioned above, these technologies can help all individuals and it can help societal resilience.

- **Resilience/defensive/offensive**
  The idea can be used in a defensive and offensive way to counter hybrid threats.

BOX 8 ENABLING TECHNOLOGY

- **Which technologies are critical in fielding the idea?**
  The databases which are used for the proven facts needs to be verified all the time and needs to be certified.
  Faked browser plugins could be produced and offered which doing false verification checks.

BOX 9 Implementation

- **Are there any restrictions with respect to using the solutions, e.g.: legal, ethical, security, etc.?**
  No.

BOX 10 Implementation effort

- **Indication of costs**:
  Describe the types of efforts and costs needed to implement the idea.
  1-2 Mio €/a until the base system is ready

  after some regular costs for maintenance and improvement

BOX 11 COUNTERMEASURES

- **Are there any potential countermeasures that could degrade the effectiveness of the solution**?
  Wrong verified information stored in the database lead to false positive checked news.

- **How durable is the idea (how long is the idea expected to be effective/useful?)**

	The solution is expected to be useful for a very long time, with a steady maintenance and efforts for monitoring the newly added or changed entries to many services of social media, news platforms and maybe messengers.
BOX 12 Preconditions (optional) - **Have all preconditions been met for the idea to be ready for implementation?** No preconditions are needed. Current browsers are supporting all necessary techniques for this.	**BOX 13 Life cycle maintenance (optional)** - **Describe who will operate, maintain, update, and upgrade the described idea.** The services shall be offered by more than one operator. Maybe founded installations and/or commercial instances from e.g. news agencies who are interested in the correctness of their reports, pictures/videos or audio streams.
BOX 14 MISCELLANEOUS **Any additional remarks/disclaimers/comments/information you might want to provide**	

5. CONCLUSIONS

5.1 SUMMARY

This Deliverable presents the work completed during the 3rd cycle of the EU-Hybnet H2020 project, and more specifically under Task 3.2 titled 'Technology and Innovations Watch'. This Deliverable has served to provide ideas and innovations for countering different dimensions of Hybrid Threats.

For the **first Core Theme, Future Trends of Hybrid Threats**, and more specifically the *primary context/critical threat relevant to political failure*, the idea presented is a mobile application to pinpoint acts of harassment & violence on the street and online. While the main beneficiaries are the law enforcement and rescue agencies, the idea also engages the public in building resilience against attacks on societal coherence. For the *primary context of new agit-prop*, the idea presented is an anti agit-prop and hostile conspiracy warning platform. The idea is based on the European Commission's Action Plan against Disinformation, and the main outcome expected is better situational awareness about the spread of designed and targeted disinformation, agit-prop and conspiracy theories. For the *primary context of alternative reality*, the idea suggested is WeVerify, a video plugin to debunk fake news on social media that spread conspiracy theories, with the aim of helping to respond quickly and effectively to the spread of disinformation. Furthermore, the EU funded EXPERIENCE project is recommended, as it is developing the technology required to help users easily create their unique VR environments, significantly improving their virtual experiences. In the future, this can be used to treat very common pathologies, such as depression, anxiety and stress, which are commonly a cause of isolation, which, in turn, lead to detrimental effects on societal cohesion.

Regarding the second Core Theme, Cyber and Future Technologies, for the *primary context of stealing data/attacking individuals*, an important primary context studies is doxing, which involves publicly exposing someone's real name, address, job, or other identifying info without a victim's consent, aiming to humiliate, bully, harass, or otherwise harm a victim. The solution offered is commercially available software that helps protect personal information against data loss, leaks, breaches and collection by third parties. Furthermore, the software automatically scans the dark web for personal information that may have been part of a data leak or data breach and helps protect the user's personal information and avoid identity theft. Besides individuals, these solutions can help fact checkers and activists of Civil Society Organisations that are often victims of such attacks. Additionally, another software category is proposed to prevent the leak of hospital patient data, by providing secure remote access, implementing access control, encrypting data, ensuring compliance to standards in Cloud environments, providing multi-factor authentication, as well as activity monitoring and visibility.

For social engineering, software solutions are recommended that detect and stop fraud in real time with machine learning, notify customers regarding which users, devices and accounts are trustworthy, and block invalid traffic. Also, coordinated fraud attacks can be detected and suspicious patterns can be identified in real time with the Hybrid AI engine that combines neural networks with symbolic AI.

For the *primary context of online manipulation/attacking democracy* the Code of Practice on Disinformation is proposed, for which relevant players in the industry agreed for the first time in 2018 on self-regulatory standards to fight disinformation, and signed the revised Code in June 2022. Virtual Operations Support (VOS) Teams (VOST) are also recommended. The Starlight H2020 project is also

recommended, as several organisations are developing different tooling enabling the detection of misleading aspects of information.

For the *primary context of attacking on services* AI and machine learning technologies are proposed for combating distributed denial of service (DDoS) attacks, and more specifically in prevention and mitigation across different areas, like Anomaly Detection, Traffic Classification, Rate Limiting and Traffic Shaping, Behavioral Analysis, User and Device Authentication, Dynamic Network Configuration or IoT Device Security. In addition, another proposal is related to advanced surveillance systems with perimeter security.

For **the third Core Theme , Resilient Civilians, Local Level And Administration**, *the first primary context studied is related to spreading of violence.* For online political harassment and SLAPP, a network of financial and legal support is proposed, as in the case of the Foreign Policy Centre, the Justice for Journalists Foundation and the International Bar Association's Human Rights Institute, that jointly organize the European Anti-SLAPP conference. A list of resources is available at the conference webpage, including means to acquire legal and financial support. Such a network can be used to stop the intimidation of journalists which forces them to spend an enormous amount of resources and energy and prevents them from offering their service to democracy, which has a profound impact on media freedom.

For the *second primary context relative to the attack on social structure*, the Offline Face Secure Access solution is proposed to safeguard higher educational institutions from external interference, in very special cases like the penetration of sensitive facilities and highly confidential projects by foreign powers. Also, the PASID system is proposed for very high security cases and only for individuals with top access clearance to sensitive information. The Passive Authentication for Secure Authentication is based on behavioural analytics engines that utilize machine learning to transform the user's activity data into behavioural biometrics. However, it is imperative to deliberate on the legal and ethical implications associated with the collection of biometric data and acquire consent of all parties prior to the implementation of PASID solution in operational contexts.

In the context of hybrid threats, trust in healthcare systems and proper functioning of service provision during crisis periods is essential. Additionally, AI-enhanced Disaster Emergency Communications are proposed as a solution to protect the social infrastructure from potential attacks and increase the resilience of the health sector during a crisis situation. In addition, an AI-powered information sharing platform for European law enforcement authorities (LEA) investigations on child sexual exploitation and abuse material (CSEM) is also proposed, a solution developed in the frame of the GRACE EU-funded project. Also, with respect to hostile messaging delivered by private messaging apps, the antidote proposed is sharing links to already existing, freely available games.

With respect to **the forth Core Theme, Information And Strategic Communications,** and the *first primary context, media conundrum,* the Media Pluralism Monitor (MPM) developed by the European University Institute is proposed. This tool assesses the potential weaknesses in national media systems that may hinder media pluralism. The solution can be used to help mitigate the loss of market shares from quality journalistic media, by bridging the gap between true quality journalistic competitiveness (i.e.journalistic integrity), as opposed to competition in terms of which outlet will first transmit the story.

For the *second primary context, sectarianism*, a prebunking game platform named 'bad news' is proposed, in which players take the role of fake news producer and learn to identify and mimic six misinformation techniques. This gaming approach significantly improves people's ability to spot misinformation techniques compared to a gamified control group and increases people's level of confidence in their own judgements.

For the *third primary context, attack on information*, a real time fact checking browser extension is proposed, that encourages critical thinking, keeps the user well informed and most importantly, ensures that truth prevails in the digital world. Furthermore, blockchain based verification is proposed to detect and combat the proliferation of misleading images. It should be highlighted that, in order to fully address the challenges associated with visual misinformation, strong collaboration would be required between technology companies, policymakers and an inclusive society.

As detailed in the methodology chapter, each category has been assigned to the members of the consortium that study and work on these fields, and their suggestions have been thoroughly discussed in weekly meetings, while comments from the other members of the team have been taken under consideration and addressed.

The future trends core theme reveals the expectations for the future form of hybrid threats, although the disruptive nature of several technologies can influence the magnitude and impact of the attack. Furthermore, in the deliverable relevant to the short list of gaps and needs, a study on the social dimensions can improve our foresight and our understanding of the future dimensions.

5.2 FUTURE WORK

This deliverable has summarised the results of Task 3.2 "Technology and Innovations Watch" of the H2020 funded EU-HYBNET project for the project's third cycle.

As detailed in the previous cycles' work, a hybrid threat is considered to be multidimensional and time-dependent (dynamic). Therefore, in order to produce one holistic solution, specific patterns would be needed, to attribute hybrid threats timely across all domains.

The present work will be used by other tasks of the EU-HYBNET project, and more specifically in Work Package 3/T3.1, "Definition of Target Areas for Improvements and Innovations", WorkPackage2 "Gaps and Needs of European Actors against Hybrid Threats"/ T2.3 "Training and Exercises Scenario Development" and T2.4 "Training and Exercises for Needs and Gap. Also, it will support the work of WP4 "Recommendations for Innovations Uptake and Standardization" on mapping on the EU procurement landscape, creation of strategy for innovation uptake and industrialization, and compiling recommendations for standardization.

The technology and innovations mapping to specific gaps and needs of the European Practitioners, as defined for each project cycle by WP2, will be continued for the last cycle as well.

ANNEX I. GLOSSARY AND ACRONYMS

Table 2: Glossary and Acronyms

Term	Definition / Description
EU-HYBNET	H2020 project titled: Empowering a Pan-European Network to Counter Hybrid Threats
Big Data	Big data is the term used for large amounts of data collected from areas such as the Internet, mobile communications, the financial industry and healthcare, that are stored, processed and evaluated using special solutions. Therefore, usually a program is used to detect rules or anomalies within the data. [28]
Artificial Intelligence (AI)	The exact definition of artificial intelligence (AI) can vary depending on the applied area. In general, AI describes systems that are able to think or act like a human being. Some of the main skills for a system to be considered as intelligent are machine learning, natural language processing, knowledge representation and automated reasoning. An artificial intelligence learns from input data and applies the extracted rules to similar situations. By receiving feedback, it improves itself. [29]
Machine Learning (ML)	The core of most training algorithms is machine learning: based on the training data the program extracts rules that it applies to similar data in order to classify it or to react with a fitting output. Depending on the received feedback, it adjusts the rules and improves its results.[16]
Blockchain	A blockchain consists of data sets which are composed of a chain of data packages (blocks) where a block comprises multiple transactions (Nofer et al, 2017)
CEN	The European Committee for Standardization
EU	European Union
EC	European Commission
EU MS	European Union Member States
H2020	Horizon 2020
GA	Grant Agreement
DoA	Description of Action
WP	Work Package
T	Task
OB	Objective
KPI	Key Performance Indicator
IA	Innovation Arena
Satways	Satways Ltd
ZITiS	Central Office for Information Technology in the Security Sector

[28] De Mauro, A., Greco, M., and Grimaldi, M., "A formal definition of Big Data based on its essential features" Library Review 65(3), 2016.

[29] Kok, J. N., Boers, E. J. W., Kosters, W.A, van der Putten, P., and Poel, M., "Artificial Intelligence: Definition, Trends, Techniques, and Cases." *Artificial intelligence* Volume 1, 1-20, 2009.

JRC	European Commission Joint Research Centre
KEMEA	Center for Security Studies
ICDS	International Centre for Defence and Security
L3CE	Lithuanian Cybercrime Center of Excellence for Training Research & Education
TNO	Nederlandse Organisatie voor toegepast-natuurwetenschappelijk Onderzoek
LAUREA	Laurea University of Applied Sciences Ltd
HYBRID CoE	European Centre of Excellence for Countering Hybrid Threats

ANNEX II. REFERENCES

[1] Joint Communication to the European Parliament and the Council, Joint Framework on Countering Hybrid Threats, European Commission, 2016.

[2] EU-Hybnet Description of Action, Coordination and Support Action, Grant Agreement No 883054.

[3] Cullen, P., Juola, C., Karagiannis, G., Kivisoo, K., Normark, M., Rácz, A., Schmid, J. and Schroefl, J., The landscape of Hybrid Threats: A Conceptual Model (Public Version), Giannopoulos, G., Smith, H. and Theocharidou, M. editor(s), EUR 30585 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-56943-5, doi:10.2760/419776, JRC123305.

[4] European Commission, MEMO - Frequently asked questions on Regulation (EU) 2019/452 establishing a framework for the screening of foreign direct investments into the Union, 09 October 2020.

[5] JOIN(2018) 36 Final. Brussels, December 5th, 2018.

[6] Esteban, Joan-Maria, Debraj Ray, «On the measurement of Polarization", Econometrica, Vol.62, No 4, 819-851, July 1994.

[7] Blechschmidt, B., Guide to doxing: Tracking identities across the web, Blog Article, November 2014.

[8] Kentucky General Assembly, 2021 Regular Session, Senate Bills, Senate Bill 267, webpage last accessed September 2023

[9] 'What Is Doxxing, Is Doxxing Illegal, and How Do You Prevent or Report It?', Avast, Academy, online article, accessed June 2023.

[10] Shivananghan, M., Modern Engineering Explained -10 Types of Social Engineering Cyberattacks, FreeCodeCamp, March 21st, 2023.

[11] Tanant, F., Fraud Detection with Machine Learning & AI, Seon company online article, accessed September 2023.

[12] The Rise of AI-powered Fraud Detection in Payments: Securing your transactions, Sweep, May 24th, 2023.

[13] AI improves fraud detection, prediction and prevention, IBM Watson Studio, online article assessed July 2023.

[14] Invalid Traffic – What Is It and How to Prevent It?, SETTUPAD blog, June 2022.

[15] European Commission, Shaping Europe's digital future, Major online platforms report on first six months under the new Code of Practice on Disinformation, online article, published September 2023

[16] European Commission, Shaping Europe's digital future, Code of Practice on Disinformation : new reports available in the Transparency Centre, online article published September 2023.

[17] Mittal, M., Kumar, K. and Behal, S., Deep learning approaches for detecting DDoS attacks: a systematic review | SpringerLink, Soft Computing, 27, 13039–13075, 2023.

[18] Veranyurt, O., Usage of Artificial Intelligence in DOS/DDOS Attack Detection, International Journal of Basic and Clinical Studies, 8(1): 23-36, 2019.

[19] Alter, S., Violence on television, Law and Government Division, publications of the Government of Canada, October 1997.

[20] Canadian Radio-Television and Telecommunications Commission, Government of Canada, assessed July 2023.

[21] CRTC Policy on Violence in Television Programming, 1996-36, Canadian Radio-Television and Telecommunications Commission webpage, assessed July 2023.

[22] European Commission, The 2023 CHAR-LITI Call for proposals under the CERV, January 26th 2023.

[23] European Commission Expert Group against SLAPP, Minutes of Meeting, 21 November 2022.

[24] Critical and Emerging Technologies List Update , Executive Office of the President of the United States, a report by the Fast Track Action Subcommittee on Ctitical and Emeging Technologies of the National Science and Technology Council, February 2022.

[25] McGuire, W. J., Resistance to Persuasion conferred by Active and Passive prior Refutation of the Same and Alternative Counterarguments, The Journal of Abnormal and Social Psychology, 63(2), 326–332, 1961.

[26] Basol, M., Roozenbeek, J., and Van der Linden, S., "Good News About Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News," Journal of Cognition 3, no. 1, 2020.

[27] Roozenbeek, J., & van der Linden, S., Fake News Game confers Psychological Resistance against online Misinformation, Nature Palgrave Communications, 5(65), 2019.

[28] De Mauro, A., Greco, M., and Grimaldi, M., "A formal definition of Big Data based on its essential features" Library Review 65(3), 2016.

[29] Kok, J. N., Boers, E. J. W., Kosters, W.A, van der Putten, P., and Poel, M., "Artificial Intelligence: Definition, Trends, Techniques, and Cases." Artificial intelligence Volume 1, 1-20, 2009.