VIGILANT

VITAL INTELLIGENCE TO INVESTIGATE ILLEGAL DISINFORMATION

Deliverable 1.5 – Data Management Plan v3

Project Information
Project Number: 101073921
Project Title: VIGILANT: VITAL INTELLIGENCE TO INVESTIGATE ILLEGAL DISINFORMATION
Funding Scheme: HORIZON-CL3-2021-FCT-01
Project Start Date: November 1st 2022

Deliverable Information
Title: Data Management Plan
Work Package: WP 1 - Coordination and Management
Lead beneficiary: TCD
Due Date: 31/10/2025
Revision Number: V1.0
Authors: Brendan Spillane, Eva Power, Gary Munnelly, Annye Braca and Owen Conlan
Dissemination Level: Public
Deliverable Type: DMP (Data Management Plan)

Overview: This deliverable is Version 3 of the Data Management Plan for the VIGILANT project and acts as a framework for consortium partners to base their Data Management Plans on.

List of Changes and Additions from DMP v2 to DMP v3

Page or Section	Change	Reason
All	Spelling and grammar fixes	To improve readability
All	Change of tense to past tense	To reflect the end of the project and to improve readability for those reading the deliverable in the future
All	Added additional minor details	To improve clarity and understanding
All	Updated references to DMP v2 to DMP v3	To reflect the latest version of this document
4.3	Added the Data Risks and Mitigation Strategies section	Good practice and to explain how the project would handle potential data issues
3.6	Added a new section on external hosting for final integration and demonstration.	To explain that only the platform was transferred to external secure hosting and not any PA or sensitive data.
Fig 3-1	Added an image depicting example of synthetic data	To demonstrate the type of synthetic data that was used in VIGILANT's development and testing.
7.7	Added new Data Value and Impact section	To show that the consortium understands the value of the data produced in the project and how it can be used in the future.

Revision History

Version #	Implemented by	Revision Date	Description of changes
0.1	Eva Power	10/02/2025	Created new doc and made initial file name changes
0.2	Eva Power	11/03/2025	Made initial tense changes, reviewed content to identify areas requiring further detail.
0.3	Brendan Spillane	11/03/2025	Reviewed content, made further tense changes.
0.4	Brendan Spillane	2/08/2025	Improved spelling and grammar, added further detail, added additional images, finalised all tense changes.
0.5	Brendan Spillane	5/08/2025	Added the Data Risks and Mitigation Strategies section and table and external hosting section
0.6	Brendan Spillane	13/08/2025	Reviewed document for completeness
0.7	Brendan Spillane	26/10/2025	Final formatting

The VIGILANT project has received funding from the European Horizon Europe Programme under Grant Agreement No. 101073921. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the VIGILANT project or the European Commission. The European Commission is not liable for any use that may be made of the information contained therein.

Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Approval Procedure

Version #	Deliverable Name	Approved by	Institution	Approval Date
0.4	D1.5	Brendan Spillane	UCD	05/08/2025
0.7	D1.5	Brendan Spillane	UCD	26/10/2025

Review Procedure

Version #	Deliverable Name	Approved by	Institution	Approval Date
0.6	D1.5	Eva Power & Owen Conlan	TCD	28/08/2025

Table of Acronyms

Acronym	Definition
APC	Article Processing Charge
DMP v3	Data Management Plan version 3
EC	European Commission
EEB	External Ethics Board
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable
FIMI	Foreign Information Manipulation and Interference
FLOSS	Free/libre and Open-Source Software
GA	Grant Agreement
GDPR	General Data Protection Regulation
LED	Law Enforcement Directive
LLM	Large Language Models
PA	Police Authority
PC	Project Coordinator
PI	Principal Investigator
PMT	Project Management Team
POPD	Protection of Personal Data
RDM	Research Data Management

REC	Research Ethics Committee
TCD	Trinity College Dublin
SCSS	School of Computer Science and Statistics

Table of Contents

1 Executive Summary

This Data Management Plan v3 (DMP v3) sets out the updated processes for managing the data generated and collected during the Horizon Europe VIGILANT project. The DMP v3 is a key element of good data management. It describes the life cycle for the data that has been collected, processed and/or generated by the VIGILANT project. As per Section 4.5 of the Project Handbook, this DMP v3 covers non sensitive data collected, processed and/or generated as part of the development phase of the VIGILANT platform and does not cover the data collected, processed and/or generated by PAs using VIGILANT when it is deployed. This DMP is a 'live document' and has been updated throughout the project. Versions of this DMP include:

- Data Management Plan v1 - M6
- Data Management Plan v2 - M18
- Data Management Plan v3 - M36 (this version)

This document should be read in conjunction with D1.1 Project Handbook v2, D1.6 Mid Term Report, D2.1 Ethics Framework, D2.3 Ethics Oversight Report, D8.1 POPD Requirement No 1, and D8.2 AI Requirement No 2. This document is heavily based on D1.4 Data Management Plan v2 with some additional updates where necessary. The list of updates is contained in the Table 'List of Changes from Version 2 to Version 3 of the DMP' on page 2 of this document.

2 Introduction

This DMP v3 has been a "living document" and it complements the GDPR requirements of Privacy by Design and Privacy by Default by ensuring that data protection is incorporated from the outset of the project. This DMP v3 outlines the framework and the principles which individual partners' DMPs in the project should meet in line with the contract obligations stipulated in Article 15 of the VIGILANT GA. There are two main types of data this DMP v3 focuses on:

- Data (public and synthetic) ingested by VIGILANT during its development.
- Data produced by the project.

3 Data (Public and Synthetic) Ingested by VIGILANT During its Development

3.1 Approaches to Sourcing and Ingesting Data During Development Phase

Two approaches ('step-by step' and 'low hanging fruit') described in detail below have guided how the VIGILANT project has accessed data during its development phase. This cautious approach enabled the PMB to update this DMP v3 and consider the ethical, legal, security and privacy concerns of the project and as yet unknown concerns related to new and emerging formats and sources of disinformation.

3.1.1 A 'Step by Step' Approach

Disinformation is incredibly difficult to identify. Issues include:

- It is often subjective.
- It continuously changes form.
- It is disseminated in multiple channels.
- New forms and channels are created monthly/weekly/daily.
- Detection methods often only work for the more obvious forms.
- Originators / spreaders are often anonymous and or deliberately hide their identity.
- Individuals and groups may have multiple overlapping and interconnected accounts and groups.
- It can be organic or organised.
- It is reactive to societal issues.
- It is often designed to go undetected.
- It can usually be classified as 'dirty data'.
- It can be spread quickly or promoted to large audiences by bots.
- Advanced technology is used to quickly create large numbers of individual messages using a single narrative that appear to be made by humans.
- It morphs and transforms to get around barriers and other methods designed to stop its spread.
- Sources of disinformation are often 'hydra-like' and reappear in other locations and in multiple forms.
- Disinformation messages may be obfuscated or only understandable by a specific audience.
- Code words, dog whistles and hidden messages are regularly used.
- It is often disseminated through channels and mediums which are difficult/impossible to access.
- It is difficult for technical solutions to distinguish between disinformation and satire.

Due to these difficulties, and many others, the VIGILANT project adopted a 'step-by-step' approach to combating this problem. Step-by-step is defined as to progress gradually and carefully to achieve an end goal. The VIGILANT project DMP has been updated based on the results of the Police Authority (PA) requirements analysis (T2.5), ethics requirements (T2.1, T2.2 and T2.3), and based on new forms/mediums/channels of disinformation becoming prevalent on the Internet and social media during the lifetime of this project. It has also been updated based on the knowledge gained from undertaking WP3 and WP4,

specifically the generation of synthetic data necessary to train and evaluate the tools being integrated with VIGILANT. D1.3 Data Management Plan v1 was submitted at M6 and D1.4 Data Management Plan v2 was submitted at M18. This final version, D1.5 Data Management Plan v3 was submitted at M36.

3.1.2 A 'Low Hanging Fruit' Approach

The purpose of disinformation is to affect change in others. To affect change in the general public, thus it must be easily accessible and consumable by the public. While a lot of disinformation originates in the darknet, among private groups and closed networks (or its spread is coordinated in them), the ultimate aim of the originator, coordinator or spreader is for it to 'bubble up' to mediums and formats that are easily accessible and consumed by the general public to increase the likelihood that the disinformation will affect change in an individual or societal behaviour. Coincidentally, and based upon the results of the initial PA requirements gathering workshops conducted prior to the project beginning and the co-design workshops conducted as part of the requirements analysis phase (see D2.5 Detailed Requirements Analysis), it is these general (and most likely to affect change) forms of disinformation, that are linked to criminal activities, that the PAs are most interested in. For example, they are more interested in a disinformation narrative that may result in a violent anti-immigrant protest with several hundred people in their jurisdiction than a disinformation narrative that the government released the Covid-19 virus as a means of testing their ability to control people and limit their movement.

Consequently, the VIGILANT project has adopted a 'low hanging fruit' approach and initially targets the large volume of disinformation that is available on the public internet before attempting to target the harder-to-access disinformation (and less impactful) in closed networks or private groups. Low hanging fruit is defined as targeting easier goals before moving onto harder goals. Examples of relatively easy to access sources include the 816 disinformation websites identified by Papadogiannakis et al. (2023)[1], the 118 fake news evaluation datasets surveyed by Murayama (2002)[2] and the 27 evaluation datasets surveyed by D'Ulizia et al.[3] (2021), and the Covid-19 disinformation dataset by Kim et al. (2021).[4] The VIGILANT project has actively monitored new sources of disinformation throughout its development and has forged links with related EU projects to gain access to their datasets.

[1] Papadogiannakis, E., Papadopoulos, P., Markatos, E. P., & Kourtellis, N. (2023). *Who Funds Misinformation? A Systematic Analysis of the Ad-related Profit Routines of Fake News sites* (arXiv:2202.05079). arXiv. https://doi.org/10.48550/arXiv.2202.05079

[2] Murayama, T. (2021). *Dataset of Fake News Detection and Fact Verification: A Survey* (arXiv:2111.03299). arXiv. https://doi.org/10.48550/arXiv.2111.03299

[3] D'Ulizia, A., Caschera, M. C., Ferri, F., & Grifoni, P. (2021). Fake news detection: A survey of evaluation datasets. *PeerJ Computer Science*, *7*, e518. https://doi.org/10.7717/peerj-cs.518

[4] Kim, J., Aum, J., Lee, S., Jang, Y., Park, E., & Choi, D. (2021). FibVID: Comprehensive fake news diffusion dataset during the Covid-19 period. *Telematics and Informatics*, *64*, 101688. https://doi.org/10.1016/j.tele.2021.101688

3.2 Synthetic Data and Public Data

3.2.1 Synthetic datasets

Large Language Models (LLMs) are text generation models which have demonstrated an uncanny ability to produce coherent, human-like responses to text-based prompts from users. Their ability to generate large volumes of human-like text make them an exciting potential resource for generating training and evaluation data for low resource domains. Research by Zhuoyan et al. (2023)[5] has already shown that classification models trained entirely on synthetic data achieves accuracy levels that are close to identical models trained on real-world data, depending on the subjectivity of the content being classified.

During the development phase of VIGILANT, the project used synthetic data generated with the aid of off-the-shelf LLMs during the development of the platform. There have been both benefits and insights gained from this approach. VIGILANT has further researched into the suitability of synthetic data as a substitute for real world data when building tools to tackle mis/disinformation. Crucially, these insights have been a benefit for strengthening the ethics of gathering training data for other projects like VIGILANT. Through its work on identifying and collecting existing disinformation datasets, there is deep knowledge within the consortium of the characteristics of the phenomena and the representations that has been used to train models. Using synthetic data where possible protects the privacy of individuals on the Internet by reducing the need to scrape real data.

To date, research has focused on the use of publicly available LLMs on the popular hosting site Hugging Face.[6] As the project avails of these tools for the creation of training data, the partners are learning more about the susceptibility of various LLM to exploitation by bad actors who may seek to flood the information space with data generated en masse by virtual agents.

```
[RACIST1] And it's not just the leftists destroying real women. We've got these
 foreigners flooding into our countries, bringing their backward cultures with
them. Their women are treated like slaves, and they don't give a shit about our
 way of life. All they care about is stealing our jobs and taking advantage of
our generosity.
```

Figure 3-1 Example of right wing anti-immigrant synthetic disinformation data generated via LLM for the project.

Synthetic data has been generated based on the use cases provided to VIGILANT by PAs during the requirements analysis phase of the project (see D2.5 Detailed Requirements Analysis). These use cases help to identify the types of problems that are faced by PAs and were used to craft prompts for the LLM. The project generated synthetic data to mirror the real-world experiences of PAs described during co-design workshops undertaken as part of the detailed requirements analysis, and subsequently use this data to train, evaluate, and demonstrate VIGILANT tools. To ensure that the synthetic disinformation generated by the LLMs was an accurate representation of the disinformation that the PAs deal with, members of partner PAs reviewed the

[5] Li, Zhuoyan, et al. "Synthetic data generation with large language models for text classification: Potential and limitations." *arXiv preprint arXiv:2310.07849* (2023).
[6] Hugging Face https://huggingface.co/

generated data and confirmed that it is representative of the content that they regularly experience. As a result of this work, the project has developed a method to generate a virtually unlimited supply of synthetic disinformation data without the need to harvest personal data of Internet users. This work is now being prepared for academic publication so that other related projects can use this method.

3.2.2 Public Data

Despite the significant efforts of project partners to generate synthetic data, it must be acknowledged that this work was experimental. Hence the platform also ingested limited amounts of publicly available data to train, evaluate, and demonstrate VIGILANT tools from three main categories of sources:

- From public social media sources which have been identified as being common sources of disinformation from previous H2020 or Horizon Europe projects, from peer reviewed academic publications, or from anti-disinformation projects such as European Digital Media Observatory (EDMO)[7] or the EU DisinfoLab,[8] or those identified from PAs and other project stakeholders during co-design workshops.
- From public fact-checking websites, such as members of the International Fact-Checking Network[9] (IFCN), European Fact-Checking Standards Network[10], and public fact-checking aggregation services, such as H2020 project WeVerify's DataBase of Known Fakes[11] (DBKF) or Google's Fact Check Explorer.[12]
- From public news and fringe websites, blogs, and discussion forums such as the 816 disinformation sources identified by Papadogiannakis et al. (2023)[13] which have been identified in previous research and projects as being common sources of disinformation.

As shown in Table 1-1, due to the nature of the project, in some circumstances some public data ingested from the Internet and social media sources may contain personal information. Note that the types of disinformation shown in Table 1-1 are not mutually exclusive. Despite efforts to target known public sources of disinformation for ingestion, the project has likely ingested normal or general information (non-disinformation). Some of the many difficulties which make distinguishing between information and disinformation an exceedingly difficult problem to solve are shown in Section 3.1.1. Therefore, it is important to note that during the development phase, when some special categories of personal data have been ingested as part of the effort to ingest disinformation, very little of this has been seen by the development team or has been surfaced in analysis results unless it is related to specific test and evaluation scenarios developed for the project which were derived from example PA use-cases and workflows.

[7] The European Digital Media Observatory https://edmo.eu/
[8] EU Disinfo Lab https://www.disinfo.eu/
[9] International Fact-Checking Network https://www.poynter.org/ifcn/
[10] European Fact-Checking Standards Network https://efcsn.com/
[11] Database of Known Fakes https://weverify-demo.ontotext.com
[12] Fact Check Explorer https://toolbox.google.com/factcheck/explorer
[13] Papadogiannakis, E., Papadopoulos, P., Markatos, E. P., & Kourtellis, N. (2023). *Who Funds Misinformation? A Systematic Analysis of the Ad-related Profit Routines of Fake News sites* (arXiv:2202.05079). arXiv. https://doi.org/10.48550/arXiv.2202.05079

Table 3-1 Categories of personal data which VIGILANT may ingest.

#	Type	Likelihood / Risk	Description	Examples
1	Personal data about a celebrity or political figure which is generally known or widely available	High / Low	Information which can easily be found about a public figure, from e.g., Wikipedia or news articles.	- Hillary Clinton's family background, religious affiliation, and political beliefs. - George Takei's sexual orientation and backing of social causes and political movements. - Images of a politician's family (spread by a third-party spreading disinformation) which the viewer could use to draw conclusions about children's ages, genders, locations they visit etc.
2	Personal data about a private individual that is available publicly.	High / Medium	Information a person chooses to share on social media, blogs, or Internet forums.	- John Doe went to X restaurant on the 28th of Nov. - John Doe's opinions on his religion versus another religion, how the women in his family should behave, and on the laws politicians should pass. - Jane Doe's unfavourable comparison of her children (name, age, gender, school grades, sexual orientation etc), who are raised in a liberal or 'woke' environment) versus the children of a political figure who are raised in a religious environment. - A private individual sharing their union membership and voting history on Twitter. - A private individual who includes details of their political affiliation, beliefs and location in their social media name e.g., "RepublicanMike", 'NewYork_gay4women', 'JaneHatesScottishToryScum' or 'WhitePowerDave_to_END_Trans_indoctrination'. - Jane Doe who regularly posts pictures of herself, and details of her family, car, job and attendance at political events and conservative rallies in an anti-climate change disinformation group.
3	Personal data derived from public social media data	High / Medium	Information which can be derived from connections in publicly available social media data or derived from analysis on large amounts of public data taken together.	- The times which Jane Doe tends to post may hint at the time zone she resides in. - People Jane Doe may know in real life as she often shares their posts. - Jane Doe's is part of a group which retweets content from several source accounts with reference to her family and personal circumstances. - John Doe's membership of multiple disinformation groups (anti-vaccination, climate change denial, anti-immigrant) using different but related names derived through indicators in the text of his public posts. - Private information about an individual who is targeted by a disinformation group or campaign e.g., private details about an individual who conspiracy theorists believe is a 'crisis actor' who is working undercover for the police or government.

3.3 Deleting Data

During the development phase of VIGILANT, any publicly available data ingested has been deleted automatically after 90 days. In some limited circumstances, some data was held for up to 120 days. This was to test features which allow PAs to mark data as being pertinent to an investigation and thus it needed to be stored for longer.

In some limited circumstances during the development phase, data has been anonymised and retained by partners for up to 6-months and used in model training and evaluation for the VIGILANT Toolbox. Partners also had the option to submit a request to the Data Compliance and Monitoring Team (see section 3.5 below) to retain data for longer periods for project purposes or for the option to publish any fully anonymised non-personal datasets.

Once deployed, the period of time for which data is stored can be adjusted to comply with each PAs needs, the Law Enforcement Directive (LED), or other national laws and regulations. A data export facility has been integrated into VIGILANT to ensure that data pertinent to longer investigations can be retained by PAs on external (cold) storage which would reside outside of the scope of the VIGILANT system.

3.4 AI and Prediction

The tools and technologies which have been integrated into VIGILANT are designed to analyse data so that PAs may respond to or prevent potential criminal acts. They are not designed to reason across large datasets of private or personal data to identify or target individuals. The project is also not using any predictive models to identify individuals who may commit crimes.

The AI tools do not possess active learning capabilities so as to ensure that any potential bias present in a user of the system would not translate onto the AI tools and their performance, i.e. given the same input the AI does not produce a different output based on who the user is, or what their (biased) beliefs might be, thus ensuring that all officers have the same user experience.

3.4.1 Compliance with the new AI Act

VIGILANT's partners have been conscious that the EC has been moving towards the regulation of AI for several years (see timeline of developments[14]) and have been actively monitoring the progress of the new AI Act[15] since the EC proposed the first EU regulatory framework on AI in April 2021 and was formally adopted by the European Council on the 21st May 2024. Several actions have been taken to ensure that the project was compliant with this new regulation when it comes into force. These include:

- Ensuring that it is a regular item at Executive Board and General Assembly meetings.

[14] AI Act Timeline of Developments - https://artificialintelligenceact.eu/developments/
[15] AI Act Explorer - https://artificialintelligenceact.eu/ai-act-explorer/

- Active engagement by the PMT with the Ethics Lead.
- Organising workshops with our Ethics lead ALUF to keep partners updated.
- Preemptively suspending work on the facial recognition tool pending the final implementation of the platform.
- Consulting with the external Ethics Advisory Board when necessary.
- Undertaking an ethics assessment of tools integrated with the platform and subsequent versions of VIGILANT, see D2.2 Ethics Guidelines for PAs.

3.5 Compliance and Monitoring

A Data Compliance and Monitoring Team has been set up consisting of the PC, PEO, and TC to ensure that the VIGILANT project is adhering to all necessary regulations. Data compliance and monitoring has also been added to the agendas for the bi-monthly EB and six-monthly GA meetings. As the VIGILANT project progresses, external experts have been asked to review the project to ensure that all necessary laws and regulations are being complied with during development and deployment phases.

3.6 Final Development and Testing

During most of the VIGILANT project's development phase, the majority of core development work, including integration of platform components and tools, and hosting of sensitive data was executed within the secure computing environment at the ADAPT Centre in TCD as this environment offered a controlled and GDPR-compliant domain. Prior to transfer to TCD for integration, individual partners undertook development and testing work on their tools and components on their infrastructure. This ensured a consistent, centralised environment for final builds and testing.

In the final stages of the project, the VIGILANT platform outgrew the secure hosting capacity ADAPT, which began to negatively impact the development and testing of the platform. As a result, it was decided to host the platform on a private server with Hetzner[16] in Nuremberg, Germany, one of the largest data centre operators in Europe. Hetzner is GDPR compliant and ISO 2701 certified[17] demonstrating robust data protection and internationally recognised information security management standards. Hetzner has strong physical, cyber and operational security measures including video monitored perimeters and access control to server rooms. Only the VIGILANT platform and limited data for demonstration and evaluation was transferred to Hetzner. No PA data, sensitive training data, or partner data was transferred. When new hosting infrastructure comes online in ADAPT, the Hetzner hosting will be wiped clean and shut down.

This migration not only alleviated the operational constraints of ADAPT's infrastructure but also successfully demonstrated the VIGILANT platform's readiness for deployment in real-world, externally hosted environments.

[16] https://www.hetzner.com/
[17] https://www.iso.org/standard/27001

3.7 Data Value and Impact

The datasets and synthetic data generation capabilities developed in the VIGILANT project have significant long-term value for research by consortium partners and the public where they have been made available. However, as described in section 4.2, some of the data generated in the project cannot be shared externally due to issues related to security and potential harm. This is also explained in detail in D8.1 POPD. The consortium is committed however to making data available to trusted researchers in their networks or through trusted research networks.

3.8 Data Summary

3.8.1 Data Sourcing

As explained in section 3.2 of this deliverable and in D8.1 POPD, during the development phase of VIGILANT a limited amount of public data from the general Internet and publicly available social media sources was ingested. A range of management and security precautions have been put in place to ensure the highest standards of ethics and privacy. This is essential to build and test VIGILANT so that it may fulfil its purpose of detecting disinformation linked to criminal acts. VIGILANT adopted a 'low hanging fruit' approach to collecting publicly available disinformation. In the early stages of the project the publicly available disinformation that was collected was from easily accessible sources such as the 816 disinformation websites identified by Papadogiannakis et al. (2023)[18], or from other public sources identified during the project. As the project developed, more difficult to access public sources of disinformation were targeted for data scraping. By adopting this approach, the project was able to institute appropriate data collection and data management guidelines to suit situations as they develop.

3.8.2 Data from Partners or Focus Groups

Limited data from the PAs involved in this project was collected as part of the design and evaluation of VIGILANT. This data was mostly the result of co-design workshops and from task based or survey evaluation methods. This data was only accessible to those within the consortium. Anonymised and aggregated results and findings have been published in project deliverables and other project publications.

Limited personal data was collected as part of public events which were held as part of VIGILANT's DC&E commitments. This includes events at European Researchers Night which gathered information from the public about their opinion of projects like VIGILANT, and what type of responses should be put in place. Anonymised and aggregated results and findings were published in project deliverables and peer reviewed journals and conference proceedings.

[18] Papadogiannakis, E., Papadopoulos, P., Markatos, E. P., & Kourtellis, N. (2023). *Who Funds Misinformation? A Systematic Analysis of the Ad-related Profit Routines of Fake News sites* (arXiv:2202.05079). arXiv. https://doi.org/10.48550/arXiv.2202.05079

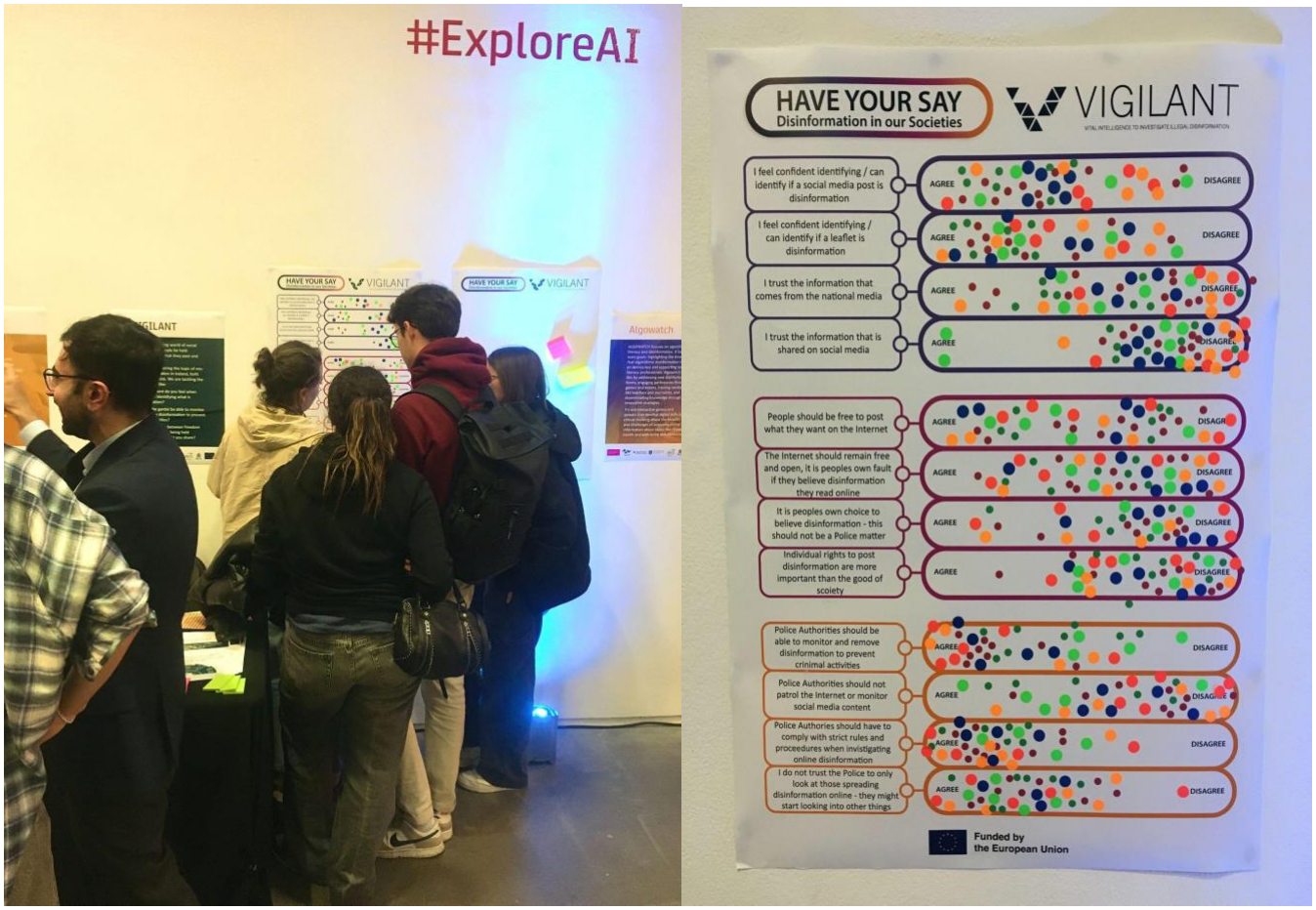

Fig 3.1 Images showing engagement at European Researchers Night event held in TCD.

3.8.3 Data Processing

Due to the nature of the personal sensitive data that was obtained during project development, as per requisition of the European Commission (EC), the consortium created a Protection of Personal Data (POPD) deliverable (See Deliverable 8.1) that specified the measures that were followed at all stages of data handling during project. Among the key guidelines followed as set out by the POPD, was the “data minimisation principle”. This ensured that any personal data obtained during the project was processed only in cases in which it was relevant to the project purposes and only while it was necessary to fulfil these purposes.

Due to the sensitive nature of some data that was ingested (examples shown in Table 1-1), the project consulted with its External Ethics Board (EEB) (see Section 1.6.3 of the D1.1 Project Handbook) about the use of this data to ensure that adequate protections are in place at all times.

3.8.4 Data Utility and Storage

During the development stages of the project data:

- The majority of VIGILANT data has been stored and processed on secure storage provided by the ADAPT Centre in TCD and was only accessible to VIGILANT partners. The ADAPT Centre[19] based in TCD is a €100m Science Foundation Ireland National Research[20] Centre for AI Driven Digital Content Technology. It has extensive secure hosting infrastructure integrated within TCDs secure hosting and has been involved with 63 H2020 and Horizon Europe projects including those storing and processing sensitive personal health data.
- Limited subsets of data have been stored on individual technical partners for the purposes of developing tools. VIGILANT partners have significant experience in secure data handling and protection from their day-to-day business, research and academic operations and from their involvement in FP7, H2020 and Horizon Europe projects. As a result, each partner already has operational policies in place regarding potential ethics issues as well as privacy and security guidelines for data protection, adhering to national and EU regulations. Ultimately, each partner is responsible for their servers' data protection and security mechanisms.

All data has been treated as confidential and accessible only to those involved in the VIGILANT project. Only partners from TCD had access to VIGILANT data on ADAPT's secure servers. In accordance with the data minimisation principle, data was only stored while it was relevant, so that data storage was limited to the purposes of the research project. As mentioned above in Section 3.3, during the development phase, data was deleted automatically after 90 days. In some limited circumstances, some data was held for up to 120 days to evaluate data sustainment capabilities being developed for PAs.

3.8.5 Data Sharing

Data was only shared between VIGILANT partners from countries which have adopted GDPR 2016/679 regulations or their equivalent.

All instances in which data is shared with Moldova and the United Kingdom, as non-EU countries, demand special attention to ensure that the transfer of such personal data from the EU to them took place in accordance with the abovementioned legislation, namely Chapter V of the General Data Protection Regulation 2016/679 and the EU Directive 2016/680. Moreover, the project ensured that personal data is processed in the EU except in some rare and limited circumstances where it has been processed by USFD in the UK in which case EU rules and regulations such as GDPR were applied. As per D8.1 POPD Requirement 1, in some limited circumstances, KInIT may transfer some personal data to the United Kingdom and the United States for project activities during the development phase. This transfer of data is in accordance with Chapter V of the General Data Protection Regulation 2016/679 and with the EU Directive 2016/680 and appropriate safeguards have been put in place.

The VIGILANT project has been actively engaging with other stakeholder organisations who have a mandate from the EC, from national European governments, friendly partner nations, and organisations who are actively involved in detecting, analysing and combating disinformation linked to criminal activities, Foreign Information Manipulation and Interference (FIMI), and or

[19] https://www.adaptcentre.ie/about/
[20] https://www.sfi.ie/sfi-research-centres/

other forms of harmful content and threats. These include other funded Horizon Europe projects (e.g., FERMI,[21] REGROUP,[22] ARM[23], ATHENA,[24] EU-HYBNET[25]), Europol Innovation Lab[26], the European Anti-Cybercrime Technology Development Association[27] (EACTDA), The European Centre of Excellence for Countering Hybrid Threats[28] (Hybrid CoE), and the CyberPeace Institute.[29] As these projects are all closely associated with the EU or with individual member states and because their focus is also on combatting disinformation, the VIGILANT project may decide to share some project data with them from time to time.

3.8.6 Data Protection and Security

During the development stages of the project, the VIGILANT platform and data was hosted on secure servers in the ADAPT Centre; a Science Foundation Ireland national research centre primarily located in TCD. ADAPT hosts dozens of other EU projects including some with very large health and disease datasets. Data was regularly deleted after 90 or 120 days. Additional security measures followed during the project development phase include:

- Heavily restricted access control (system authorisation) to ensure data was accessible to cleared members of the consortium through a VIGILANT personalised login. Once deployed on PAs own systems, it was only available to designated Officers logged into the PAs network and behind their firewalls and other security processes.
- User control: Data could not be accessed, copied, altered nor deleted without authorisation and clearance and all data activity was strictly logged and recorded and subject to auditing.
- Separation control: Collected data is processed separately by 1) setting different access rules for study datasets according to the purpose they serve and 2) having a network of federated repositories and processing resources.
- Communication control: Data shared between participants should occur only to achieve project goals and/or tasks that require cooperation and collaboration between members of the consortium and in accordance with the data minimisation approach.
- Transmission control: Unauthorised reading, copying, alteration or deletion in data transfers is prevented by using SSL/TLS protocols to implement authenticated encrypted communications. To ensure formattable files are not edited, these were saved as read-only to avoid accidental changes or overwriting.
- Input control: VIGILANT is designed to only ingest publicly available data posted in e.g., public Facebook groups or from public YouTube and Twitter accounts etc., not to track personal accounts or private information on any platform.

For more detailed information on our data protection security guidelines, please refer to our D8.1 POPD deliverable.

[21] Fake News Risk Mitigator (FERMI) - https://fighting-fake-news.eu/
[22] Rebuilding Governance and Resilience Out Of The Pandemic – REGROUP - https://regroup-horizon.eu/
[23] The Long Arms of Authoritarian States (ARM) - https://cordis.europa.eu/project/id/101132437
[24] An exposition on THe forEign informatioN mAnipulation and interference (ATHENA) - https://project-athena.eu/
[25] European Union Horizon 2020 Project EU-HYBNET (https://euhybnet.eu/) GA #883054
[26] Europol Innovation Lab - https://www.europol.europa.eu/operations-services-and-innovation/innovation-lab
[27] EACTDA - https://eactda.eu/index.html
[28] Hybrid CoE - https://www.hybridcoe.fi/
[29] CyberPeace Institute https://cyberpeaceinstitute.org/

3.8.7 Ethical Aspects

The VIGILANT project balanced the needs of PAs with an ethical-by-design approach and as such it has significant ethical oversight and contributions. These include:

- Ethics Assessment Framework (T2.1 > D2.1) to enable the consortium members to develop the system in accordance with the most up-to-date ethics standards and to ensure that the developed platform has the highest possible positive ethical impact with minimal ethical risks.
- Ethics Guidelines for PAs (T2.2 > D2.2) which provide guidance to decision-makers in PAs for the adoption of VIGILANT.
- Ethical Oversight (T2.3 > D2.3) which includes the setting up of an Ethics Advisory Board by M3 which included members external to the VIGILANT consortium to provide independent oversight, advice, and support to all consortium members to fulfil ethics requirements, (see External Ethics Board).
- A POPD (Protection of Personal Data) deliverable (D8.1) justifying the processing of sensitive personal data following the data minimisation principle, establishing the security measures implemented to prevent unauthorised access to such data, and ensuring all data transfers between an EU country to a non-EU countries occur in accordance with the 2016/679 GDPR and with EU Directive 2016/680.
- An AI human right impact assessment (D8.2) providing detailed information on how the respect for fundamental human rights and freedoms were ensured during the project.
- Legal and Security Compliance Plan to set out the legal and security guidelines followed by project members.

When focus groups our public engagement events such as attendance at European Researcher Night events (see section 3.8.2 and D7.7 DC&E and Training) or similar are being conducted, participants were required to sign informed consent forms which made clear the purpose of the event, what data was collected and recorded, how it was used, and for how long it was kept.

All VIGILANT research activities and data collection respect fundamental ethics principles, including those reflected in the Charter of Fundamental Rights of the European Union (EU). The Coordinator (TCD) has clear institutional guidelines on ethical issues and the TCD School of Computer Science and Statistics (SCSS) Research Ethics Committee (REC) operates in accordance with relevant national and EU legislation. Relevant EU legislation includes:

- General Data Protection Regulation[30] (EU 2016/679) on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.
- European Union Directive 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal

[30] https://eur-lex.europa.eu/eli/reg/2016/679/oj

offences or the execution of criminal penalties, and on the free movement of such data - also known as the Law Enforcement Directive.

- Council of Europe Recommendation[31] 83/10 on the protection of personal data used for scientific research and statistics.
- Council of Europe Recommendation[32] 97/18 on the protection of personal data collected and processed for statistical purposes.
- TCD Policy of Good Research Practice.[33]

4 Data Produced by the Project

This section describes in detail how data is collected, managed and processed by the VIGILANT project.

4.1 Data that has not been Published

Due to the nature of the VIGILANT project, much of the data produced to develop and evaluate VIGILANT was not made publicly available. This includes:

- When it contravened section 13.1 of the Grant Agreement (Sensitive Information). In total, eleven deliverables from this project are classified as sensitive.
- Any PA data or datasets including data collected during co-design workshops or other project activities that could identify an individual from a PA.
- Data which could result in public harm.
- Data which included extremist or violent content which should not be made available to the public.
- Data which could show nefarious actors how to produce disinformation or other forms of harmful content.

The following sections and questions/answers specifically refer to when data is being made publicly available. They do not apply to any of the above listed situations.

4.2 FAIR Data Principles

When data was being shared publicly, it was made Findable, Accessible, Interoperable and Reusable (FAIR). The answers to the following sections were guided by the H2020 Programme Guidelines on FAIR Data Management[34]. The individual questions

[31] https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016804bc647
[32] https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680508d7e
[33] https://www.tcd.ie/about/policies/Good_Research_Practice_June2021.pdf
[34] H2020 Programme Guidelines on FAIR Data Management in Horizon 2020 projects version 3.0 https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

were taken from the H2020 Data Management Plan template[35]. All answers are general answers for the consortium and project. Where appropriate, further information has been provided for certain specific instances or examples to clarify or when it is necessary to explain exceptions.

4.2.1 Data Collection in VIGILANT

- **What was the purpose of the data collection/generation and its relation to the objectives of the project?**

  The purpose of generating and collecting data was to help develop and evaluate the VIGILANT platform and its individual advanced disinformation detection and analysis components. This was required to meet the key objectives of the project, in particular KO1, KO2, K03 and K04.

- **What types and formats of data did the project generate/collect?**

  Data was generated and collected in a variety of formats including .csv, .json, .docx, .xls .txt, .xml, .pdf, and .html.

- **Will data be reused and how?**

  Where possible, the project endeavoured to use synthetic data (see Section 3.2.1), however, where this was not possible it occasionally reused existing public datasets rather than collect new ones. The use of existing data is described in detail in Section 3.

- **What was the origin of the data?**

  Data from different origins was used in the VIGILANT project. They include data from:

  - Synthetic data sources such as LLMs (see Section 3.2.1)
  - Other previous EU projects such as the H2020 PROVENANCE[36] and WeVerify[37] projects
  - Data collected from the public internet and social media channels (see Section 3.2.2)
  - Data generated by project activities
  - Data from published datasets and open access repositories (see Section 3.1.2)

- **What was the expected size of the data?**

  During the development stage of the project, several terabytes of data was used or generated by the project.

- **To whom might it be useful ('data utility')?**

  As much of the project activities are classified as sensitive, the majority of data collected for the purposes of building and evaluating VIGILANT was only available and therefore useful to members of the consortium.

4.2.2 Making Data Findable

In some limited circumstances, datasets which have been created for project development purposes e.g., for the training of a specific tool or for evaluation purposes, and which do not contain any PA or sensitive personal data, will be made available on the VIGILANT website (https://www.vigilantproject.eu/) at the end of the project or in the months thereafter. The website will be kept live for three years post project. The VIGILANT website also includes all non-sensitive VIGILANT deliverables,

[35] H2020 DMP Template https://ec.europa.eu/research/participants/data/ref/h2020/gm/reporting/h2020-tpl-oa-data-mgt-plan_en.docx
[36] 3.2.2 - Providing Verification Assistance for New Content - https://cordis.europa.eu/project/id/825227
[37] Wider And Enhanced Verification For You https://cordis.europa.eu/project/id/825297

publications, and other project material. Publications and datasets were made available on open access repositories (e.g., CLARIN Centres, Zenodo, arXiv, GitHub, TCD's TARA, or UCD's Research Repository). More datasets may be made available at the end of the project or in the months after it finishes. Where appropriate, datasets are marked up with machine readable metadata.

- **Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?**
  All datasets which have been made publicly available have been made discoverable with metadata and standard ID number. This practice will continue for any datasets released after the project is finished.
- **What naming conventions do you follow?**
  All VIGILANT project datasets have used descriptive names which use the following format 'VIGILANT_DatasetName_Task/Domain/Purpose_Number' e.g. VIGILANT_DisinfoConversation_EventAnalysis_v2. Individual datasets released by partners relating to a specific tool or task have used their naming conventions or follow those of the publication venue.
- **Will search keywords be provided that optimize possibilities for reuse?**
  Yes, these have been provided when publishing the datasets. All datasets also have a ReadMe file with such information.
- **Do you provide clear version numbers?**
  Yes, all datasets use numbering in the form of v1, v2, v3 etc.
- **What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.**
  Title, description, keywords, creator information, project information, data or creation and publication, version number, access rights, accompanying publication, organisation who created it, DOI or URL, data types, formats, sizes, access rights, data collection / generation methods, licence, citation, usage notes, software required, source, version history and changes, documentation location, other metadata as required depending on dataset, purpose and publication norms.

4.2.3 Making Data Openly Accessible

Where possible, the VIGILANT has been committed to making datasets openly available to the public and researchers in the community. However, as described above, some project datasets have not been released publicly or have only been made available to established researchers in the domain who are working to combat disinformation and other related forms of harmful content. Descriptions of datasets that cannot be made publicly available have been published along with contact information to make them findable so that access may be requested.

- **Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.**

  As described above, the intention of the project was that all project data that was not classified as sensitive, which does not contain any personal data, which does not contain PA data (or data from PA design workshops or evaluations), or which does not harm the public was made openly accessible. This includes datasets collected as part of the development of VIGILANT and for training specific tools.
- **How will the data be made accessible (e.g. by deposition in a repository)?**

  Using open access repositories such as Zenodo, arXiv, CLARIN Centres, GitHub, TCD's TARA, UCD's Research Repository or other institutions open access repositories.
- **What methods or software tools are needed to access the data?**

  All datasets made publicly available are accessible by open source or standard tools and software commonly used by practitioners in the domain or easily accessible to any computer scientist or engineer. No closed format datasets requiring specialised software or tools have been created or published.
- **Is documentation about the software needed to access the data included?**

  Where necessary, it has been included in the ReadMe files in the metadata.
- **Is it possible to include the relevant software (e.g. in open source code)?**

  No, this was not possible or necessary
- **Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.**

  Open access repositories such as Zenodo, arXiv, CLARIN Centres, GitHub, TCD's TARA, UCD's Research Repository or other institutions open access repositories.
- **Have you explored appropriate arrangements with the identified repository?**

  All technical partners in the project are already familiar with these, and most have already used them. Some datasets and source code are already available on them (see list of project publication in Section 9.4 D1.6 Mid Term Report).
- **If there are restrictions on use, how will access be provided?**

  No licence restrictions are envisioned. As explained at the start of this section, some datasets are not made publicly available due to the possibility that they could result in public harm, may include harmful or extremist content, or which could inform the development of methods or tools which could be used to create disinformation. The project and partners have and will continue to make this material available to known researchers in the field on request after the project is finished.
- **Is there a need for a data access committee?**

  As described in Section 3.5, a Data Compliance and Monitoring Team has been set up consisting of the PC, PEO, and TC to ensure that the VIGILANT project is adhering to all necessary regulations. This team has advised and made decisions on data access.
- **Are there well described conditions for access (i.e. a machine readable license)?**

  As datasets are published, all access conditions have been described in the metadata and in easily accessible formats.

- **How will the identity of the person accessing the data be ascertained?**

  In most instances, this has not been recorded. However, some datasets which may contain harmful content (see above) have only been released to established researchers in the domain. Descriptions of such datasets have been published along with contact information to make it findable so that access can be requested.

4.2.4 Making Data Interoperable

The VIGILANT project is committed to making publicly available datasets interoperable. Efforts have included using standard naming conventions, data forms, vocabularies, and qualified reference.

- **Are the data produced in the project interoperable, that is allowing data exchange and reuse between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?**

  Yes.
- **What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?**

  The project used the Dublin Core Metadata Initiative[38] when publishing publicly available datasets.
- **Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?**

  Where possible, yes. Where this is not possible, the project has provided mappings to common ontologies.

4.2.5 Making Data Reusable

All datasets included descriptive metadata and Readme.txt files to make it easier for other researchers to reuse them. Each dataset includes a clear and accessible data usage licence stipulating how and where the data can be reused.

- **How will the data be licensed to permit the widest reuse possible?**

  Where possible, all project datasets were made available using the Creative Commons Attribution (CC BY) or Open Data Commons Attribution (ODC-By) licenses. However, some partners may have specific licence requirements for datasets which they create as part of the development of individual tools.
- **When will the data be made available for reuse? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.**

[38] https://www.dublincore.org/

Most datasets that have been made public have been made available without an embargo. However, some publication venues require embargos when publishing datasets due to involvement in a challenge task or similar. Data has been made public as soon as possible, where possible, and when possible (see above for restrictions).

- **Are the data produced and/or used in the project usable by third parties, in particular after the end of the project? If the reuse of some data is restricted, explain why.**
  Yes, when it is made public (see above for restrictions).
- **How long is it intended that the data remains reusable?**
  There has been no time restriction on data reuse.
- **Are data quality assurance processes described?**
  Where appropriate, they have been described in the metadata.

4.2.6 Allocation of Resources

The VIGILANT consortium and its individual partners have considerable resources to ensure data access complies with the FAIR principles.

- **What are the costs for making data FAIR in your project?**
  The monetary costs of making data FAIR in the project consist of the publication costs for Open Access publications and the costs of recording, curating, formatting, and hosting of the data generated by the project. Many of the individual partners within the project also have their own personal budgets for making publications Open Access.
- **How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).**
  Some partner institutions such as TCD and UCD have offset Article Processing Charges (APC) because they participate in negotiated blanket publication agreements with the publishers. The remaining Open Access publication costs and the costs of producing and hosting recordings of VIGILANT events (workshops, webinars) are paid out of the dissemination budget shared by TCD and GLOB. There are no charges for using Zenodo, B2SHARE or other similar open access repositories. All participating institutions, as well as most journals and conferences, have rules and principles in place that require their researchers to make their data FAIR. As such, there are no extra costs to VIGILANT consortium partners for making the data FAIR. The costs, in time and effort, to upload data and publications to Zenodo or CLARIN Data centres are marginal and covered by the project and its overhead provisions.
- **Who will be responsible for data management in your project?**
  The Principal Investigators (PIs) of the project from the different institutions were responsible for data management, including making data and publications FAIR. The coordinator has overseen the implementation. VIGILANT's public deliverables, once approved by the Commission, have been made available on the VIGILANT website. Where possible, the data from partner's scientific publications has been published on GitHub (see Section 9 of D1.6 Mid Term Report). All publications and data that could be published has been stored at Zenodo, arXiv, CLARIN or B2SHARE (for models). Data that could not be published or has other restrictions has been stored in partner institutions' repositories or on

secure remote storage, at the choice of the hosting institution. Descriptions and contact information of this latter data has been published to make them findable. All the institutions participating in VIGILANT have made provisions for long term secure storage of data and publications in the form of repositories. Data that is uploaded to Zenodo and CLARIN data centres has been made available without a time limit. Data that is not open has been stored according to the rules of the owning institution. The use of these long-term repositories does not constitute a cost for the VIGILANT project.

- **Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?**

  This has been discussed at project meetings and among technical partners. Publicly available datasets have been made available in established long term repositories. The Data Compliance and Monitoring Team (see section 3.5) has the final decision over project data. Individual partners have made decisions for datasets they have created for individual tools.

4.2.7 Data Security

The VIGILANT consortium has taken extensive precautions to ensure the security of the platform and its data. This is described in detail in Section 3.6.6 of this deliverable, in Sections 5 and 6 of D1.2 Project Handbook, Section 4.1.2.6 of D1.6 Mid Term Report, and Section 9 of D8.1 POPD Requirement 1. The consortium has also established a Legal and Security Compliance Monitoring Team to oversee this aspect of the project (see Section 6.1 of D1.2 Project Handbook).

- **What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?**

  As described in Section 9 of D8.1 POPD Requirement 1 and in Section 3.4 of D1.6 Mid Term Report, and in Sections 3.6.1, 3.6.2, 3.6.3, 3.6.4, 3.6.5 and 3.6.6 of this deliverable, extensive provisions are in place to ensure secure data storage and sharing among partners.
- **Is the data safely stored in certified repositories for long term preservation and curation?**

  As sensitive data has not been released, this is not necessary.

4.2.8 Ethical Aspects

There are significant ethical considerations surrounding data and data access in the VIGILANT project. These have been discussed extensively among the PMC, the EB, among the wider consortium and with our Ethics Lead.

- **Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).**

  As described at the start of this section, data which is classified as sensitive, potentially harmful to the public good, or which could be used to create disinformation, or other forms of harmful content has not been made publicly available.

- **Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?**

  Informed consent has been included in all questionnaires. However, this data has not been made available due to other sensitivity issues described above.

4.2.9 Other Issues

The VIGILANT project is guided by Horizon Europe guidelines and international best practices when dealing with data of a sensitive nature.

- **Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?**

  No.

4.2.10 Open Access Publications

Horizon Europe requires that all publications arising from the work funded in the project are published in Open Access (green and gold route). VIGILANT has published its results as either immediate open access (open access journal), or delayed open access, for 6 months (gold route), at which point it has, or will, be deposited in a suitable publication repository as defined in the DMP v3. Partners have also or alternatively (green route) self-archive via institutional repositories and Zenodo as recommended by OpenAire[39]. Shorter reports and articles have been published without delay and made immediately available on the project's website. VIGILANT participants are also encouraged to publish their papers, findings, articles, etc., on their professional and research social network (e.g., LinkedIn, ResearchGate). The VIGILANT project should be acknowledged as follows:

"The VIGILANT project received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101073921."

The VIGILANT project has utilised:

- Open Access journals and open access repositories (e.g., Zenodo[40], TCD's TARA[41], or UCD's Research Repository[42]).
- Zenodo and CLARIN Centres[43] for data sharing.

[39] https://www.openaire.eu/
[40] https://zenodo.org/
[41] http://www.tara.tcd.ie/
[42] https://researchrepository.ucd.ie/home
[43] https://www.clarin.eu/content/overview-clarin-centres

4.3 Data Risks and Mitigation Strategies

The VIGILANT consortium recognises that the collection, processing, and storage of both synthetic and publicly available data carries a range of ethical, legal, and operational risks. Table 4-1 identifies the principal risks relevant to the project and outlines the measures taken to mitigate them. This section is in addition to and should be read in conjunction with the Risk Management Plan section 3 of D1.2 Project Handbook.

Table 4-1 Data risks and mitigation strategies.

Risk	Description	Mitigation Measures
Inadvertent ingestion of personal data	Publicly available data scraped from the internet and social media may contain personal data, including special category data, despite efforts to target disinformation sources.	Apply the data minimisation principle as outlined in D8.1 POPD; delete all ingested public data after 90 days (or 120 days in specific test cases); anonymise where possible; restrict access to authorised consortium members only.
Bias in synthetic data generation	Synthetic datasets generated using LLMs may inadvertently reflect biases in training data or misrepresent target phenomena.	Review synthetic data with partner PAs to confirm accuracy and representativeness; diversify prompt engineering; document known limitations; exclude biased or misleading outputs.
Data security breaches	Unauthorised access, loss, or corruption of stored project data.	Host data on secure ADAPT Centre infrastructure with strong access controls, encryption, and audit logging; partner servers to follow national/EU security standards; regular monitoring and penetration testing where applicable.
Loss of data integrity over time	Risk of metadata loss, file format obsolescence, or repository inaccessibility after project end.	Deposit in certified, open-access repositories (Zenodo, CLARIN) with persistent identifiers; maintain VIGILANT website for at least three years post-project; ensure partner repositories have long-term preservation policies.
Misuse of project outputs	Tools, datasets, or methods could be used to create or enhance disinformation.	Restrict access to sensitive datasets and certain tool functionalities; share methodologies responsibly; provide ethics guidance to potential users.
Non-compliance with the GDPR/LED/AI Acts	Risk of breaching EU data protection or AI governance requirements during collection, processing, or sharing.	Maintain Data Compliance and Monitoring Team oversight; conduct regular ethics reviews; consult with External Ethics Board; ensure cross-border data transfers comply with GDPR Chapter V and LED provisions.

These mitigations are applied in conjunction with the legal, ethical, and security measures described in Sections 3.6.6, 4.2.7, and 4.2.8 of this deliverable, and are monitored by the Data Compliance and Monitoring Team throughout the project lifecycle.

5 Updating this DMP v3

The PC, PEO and PM have been responsible for updating this DMP v3 and data management has been included in the agendas for the EB bi-monthly and GA six-monthly meetings. All partners have been encouraged to raise any issues with the PC, PEO, or PM that require this DMP v3 to be updated

6 Conclusion

This DMP v3 has set out the principles for Data Management in the VIGILANT project which have evolved and have been followed since the first version of the DMP. Many of these updates are the direct result of the work to produce the VIGILANT platform including the work on producing synthetic data, and interactions with partner projects and institutions and other expert stakeholders. It has addressed the public and synthetic data which have been ingested by VIGILANT as part of its development and the data which has been produced by the project. It describes the three known types of personal data which is commonly found in disinformation which may be ingested by the platform, when it is deleted, and it makes clear that no use of predictive AI has been used on such personal data. This DMP v3 also describes how the project has complied with the FAIR data principles for all data produced by the project.