FERMI

FAKE NEWS
RISK MITIGATOR

Project acronym:	FERMI
Project full title:	Fake nEws Risk Mitigator
Call identifier	HORIZON-CL3-2021-FCT-01
Start date:	01/10/2022
End date:	30/09/2025
Grant agreement no:	101073980

D3.4 FERMI Behaviour Analyses and Community Resilience Facilitators Package

Work package: 3

Version: 1.0

Deliverable type: R - Document, report

Dissemination level: Public

Official submission date: M30

Actual submission date: M36

Funded by
the European Union

Leading author(s):

Surname	First name	Beneficiary
Baronchelli	Adelaide	BIGS
Kolliarakis	Georgios	BIGS
Stuchtey	Tim	BIGS

Contributing partner(s):

Surname	First name	Beneficiary
Varela da Costa	João	INOV
Valente	Catarina	INOV
Tsakova	Gergana	UCSC - TRANSCRIME
Lo Giudice	Michael Victor	UCSC - TRANSCRIME
Vrotsou	Christina	CONV
Kokoliou	Panagiota	CONV

Peer reviewer(s):

Surname	First name	Beneficiary
Fikenscher	Sven-Eric	BPA
García	Joaquín	ATOS

Ethics reviewer:

Surname	First name	Beneficiary
Markopoulou	Dimitra	VUB

Security reviewers:

Surname	First name	Beneficiary
Bouwknegt	Léonie	Nederlandse Politie
Mattes	Tobias	BPA

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	**Modified by**
0.1	20.12.2024	Contribution planning and setting of timelines for delivery	BIGS
0.2	14.03.2025	Delivery of the Behavioural Profiler updates	CONV
0.3	28.04.2025	Delivery of the Crime estimate updates	UCSC
0.4	02.05.2025	Delivery of the Community Resilience Management Model updates	INOV
0.5	02.06.2025	Delivery of the Socio-economic Analyser updates	BIGS
0.6	09.07.2025	Integrated pre-final version	BIGS
0.7	16.07.2025/15.07.2025	Scientific/Technical Review	ATOS; BPA
0.8	25.08.2025	Security review	Security Advisory Board
0.9	05.09.2025	Ethics Review	VUB
1.0	08.09.2025	Final revised version	BIGS

Table of Contents

Executive Summary

The present report 3.4 is an advancement and recast of deliverable D3.3 from the first period of the FERMI project, which contained the basic conceptual and methodological framework, and first mock results of the threefold component of the Behavioural Analyses and Community Resilience Facilitators package. The component provides tools that can be applied to examine the consequences of crimes occurring due to Disinformation & Fake News. Online activities are expected to have an impact on crime rates, more specifically, on politically motivated crime. This report presents the analysis conducted in the 2nd period of the project and reflects upon the results and their societal and policy consequences, and, not least, their significance for further research action.

Report 3.4 has proceeded towards the factual econometric analysis of criminal statistics, augmented the evidence on the behaviour profiler by examining a second indicator, besides media literacy, the one of public trust, and updated the conceptual ramifications around the crime estimates and the Community Resilience Management Modeler as well as the next steps to be coming along the lines of obligations for said components stated in the GA.

While D3.4 is based for the sake of continuity conceptually and structurally upon report 3.3, it contains extensive new parts, to be found particularly in the following chapters and subchapters: Chapter 4, and particularly 4.2 (Model Description, and specifically Behaviour Profiler on the interconnection between Media Literacy, Trust and Disinformation, and Crime estimate), 4.2.2 on the econometric model and the criminological data analysis for Germany and Spain, 4.3 on the technical implementation and 4.4 with the results of the above analyses. Furthermore, in Chapter 5, subchapters 5.1 and 5.3 have been reworked and respectively updated.

The substantial results out of the data collection and processing in all subcomponents, which took place in the 2nd project period, is consequently reflected in the final chapter. For the sake of outlook into the future, some key takeaways from the work on this task need to be formulated. Against the backdrop of the overarching objective of FERMI, which is to promote an evidence-informed decision-making in LEAs for more effective prevention and mitigation of the detrimental societal impacts of disinformation, the complex issue of data availability, access, analysis, and usage appears to be a paramount one. The aspect of foresight and the forward-looking security policy planning for D&FN is a complementary issue to that. Furthermore, one should consider the EU-level drafting of resilience dashboards which would need an update to include dedicated indicators on hybrid threats. Not least, the concomitant challenge of which stakeholder is mandated as an agent to decide and act, both re-actively, and, even more, pro-actively in terms of anticipatory planning on countering the negative impacts of D&FN.

Abbreviations

AI:	Artificial Intelligence
AHP:	Analytic Hierarchy Process
CRMM:	Community Resilience Management Modeler
D&FN:	Disinformation and Fake News
DB:	Database
DSA:	Digital Service Act
FERMI:	Fake News Risk Mitigator
GA:	Grant Agreement
GDP:	Gross Domestic Product
GMM:	Generalized Method of Moments
HLG	High-Level Group
JRC:	Joint Research Centre
LEA:	Law Enforcement Agency
LSDV	Least-Squares Dummy-Variables
MCDA:	Multi-Criteria Decision Analysis
MCDM:	Multi-Criteria Decision Method
ML:	Machine Learning
M-MACBETH:	Measuring Attractiveness by Categorical Based Evaluation Technique
OLS:	Ordinary Least Squares
OSIS:	Open Society Institute – Sofia
PKS:	Polizeiliche Kriminalstatistik (Police Crime Statistics)
PMK:	Politisch Motivierte Kriminalität (Politically Motivated Crime)
RSF:	Reporters Without Borders
SLR:	Systematic Literature Review
UI:	User Interface

Tables

Figures

Equations

1 Introduction

The Behaviour Analyses and Community Resilience Facilitators Package focuses on three components of Work Package 3, namely the Behaviour Profiler, the Socioeconomic Analyser and the Community Resilience Management Modeler (CRMM). The present report 3.4 is the 2nd version and a substantial rework and advancement of report 3.3, which provided the description on the modelling, the technical implementation of the models and the first results obtained from their implementation. Report 3.4 has proceeded towards the factual econometric analysis of criminal statistics, augmented the evidence on the Behaviour Profiler by examining a second indicator, besides media literacy, the one of public trust, and updated the conceptual ramifications around the crime estimates and the CRMM.

While the 2nd version is based for the sake of continuity conceptually and structurally upon report 3.3, it contains extensive new parts, to be found particularly in the following chapters and subchapters: Chapter 4, and particularly 4.2 (Model Description), 4.2.2 on the econometric model and the criminological data analysis for Germany and Spain, 4.3 and 4.4 with the results of the above analyses. Furthermore, in Chapter 5, subchapters 5.1 and 5.3 have been reworked and respectively updated. The substantial results out of the data collection and processing in all subcomponents, which took place in the 2nd project period, are consequently reflected in the final chapter.

The deliverable provides a description on the modelling, the technical implementation of the models, the first results obtained from the implementation are handled as sub-components examining different aspects of the societal impact of disinformation: The former two focus on measuring the likelihood of politically motivated crimes occurring due to online disinformation, and have predominantly examined evidence from France, Sweden, Germany, Belgium, Finland and Spain, whereby criminological data for the socio-economic analysis have been drawn from Germany and Spain. The latter describes a model that analyses the connection between politically motivated crime and economic factors. In chapter Community Resilience Management Modeler the CRMM focuses on visualising these results in a simply manageable user interface (UI) as well as generating potential countermeasures for the end-users.

With regard to the Behaviour Profiler, specifically, societal-level indicators such as media literacy and trust in institutions, drawn from OECD and Eurobarometer surveys and analyses, offer additional insight into why certain countries may be more vulnerable to disinformation-driven crime. Research has shown that low trust in media and institutions correlates with higher engagement with alternative narratives, some of which may foster extremism. Conversely, societies with high media literacy and strong institutional trust tend to exhibit greater resilience to the effects of disinformation. By integrating these societal risk factors into the Behaviour Profiler & Socioeconomic Analyser, the FERMI project enhances its ability to contextualise crime likelihood beyond technical modelling, considering broader societal trends that influence disinformation propagation and its potential real-world impacts.

Throughout 2024, BIGS as the leading partner of this report, has been in intensive and extensive contact with all LEA and Police College representatives from France, Sweden, Germany, Belgium, Finland and Spain, as represented in the FERMI consortium, in order to explore and document whether the nationally collected criminological statistics are structured and disaggregated in such a way, which would allow for a meaningful econometric analysis. In practice, that should entail pre-structured datasets along categories of extremism and politically or religious motivated crime, geographic disaggregation along NUTS levels in order to allow the study of variance, and not least, temporal sequencing (weekly, monthly, etc.) in order to allow for Artificial Intelligence (AI)-enabled crime estimates to train the respective models.

Despite the extraordinary engagement and commitment of all LEA partners in the consortium, and the repeated in-depth delving into the respective national datasets, it has been a sobering fact that no datasets could fulfil the above criteria for econometric processing, but those of Germany, and to a lesser extent those of Spain. The two latter ones have been ultimately selected for showcasing the socioeconomic costs of online disinformation passing over to the offline world.

Indeed, apart from Germany, no other country in the FERMI consortium publicly reports systematic data on politically motivated crimes. However, Spain publishes data on hate crime. According to the official definition, hate crimes are defined as any criminal offense, including those committed against persons or

property, where the protected legal property is targeted because of its actual or perceived connection, sympathy, affiliation, support or membership of a group. Hate crimes are conceptually different from politically motivated crimes. However, they can convey the idea expressed by our theoretical approach to extremism and, therefore, Spain serves as the second empirical focal point to examine the Socioeconomic Analyser's cost calculation.

The model to explain the economic cost of extremist crimes via the calculation of the cost parameter, indicates how a one-unit increase in extremist crime impacts on economic welfare. The model was developed to be used for all the LEA and Police College countries in the FERMI consortium. However, since Germany and Spain represent two different use cases with different types of data available, we adapted the equation to each country.

Report 3.4 concludes by summarising the main takeaways from the Behaviour Analyses and Community Resilience Facilitators Package. It also reflects on the limitations and the challenges for future research work and policy action.

2 Aim and Significance of the Behaviour Profiler & Socioeconomic Analyser and Community Resilience Management Modeler in the FERMI project

The connection between online activities of people and impacts on the offline world stand in the centre of the FERMI project as explained elsewhere in this project (see, most recently, D3.2), the analysis of online activities opens up the possibility to predict certain forms of crimes on spatial levels. This is where the Behaviour Profiler, the Socioeconomic Analyser and the Community Resilience Management Modeler that work closely together step in. Based on sophisticated scientific and technical solutions, the Behaviour Analysis and Community Resilience Facilitators Package rests upon solid foundations to analyse the topic at hand. The aim of the Behaviour Profiler & Socioeconomic Analyser is to make the occurrence of politically motivated crime detectable (at least in terms of likelihood) and measurable (in terms of severity, measured along the lines of cost), whereas the CRMM makes said crimes explainable (in terms of risk). More specifically, not all disinformation distributed online potentially leads to crime in the offline world. This component aims to quantify likelihood and severity of crimes occurring due to disinformation, whose combined terms output a measurement of risk.

To complement this quantitative assessment, societal-level indicators such as media literacy and trust in institutions offer additional insight into why certain countries may be more vulnerable to disinformation-driven crime. Research has shown that low trust in media and institutions correlates with higher engagement with alternative narratives, some of which may foster extremism [1,2,3]. Conversely, societies with high media literacy and strong institutional trust tend to exhibit greater resilience to the effects of disinformation. By integrating these societal risk factors into the Behaviour Profiler & Socioeconomic Analyser, the FERMI project enhances its ability to contextualise crime likelihood beyond technical modelling, considering broader societal trends that influence disinformation propagation and its potential real-world impacts.

The analysis' results are fed into the Community Resilience Management Modeler through normalised impact index linked with the instance being investigated in the FERMI platform. The impact index takes in consideration the likelihood of an event occurring in the community and its potential severity, as well as socioeconomic factors offered by the Behaviour Profiler & Socioeconomic Analyser components.

This component takes a risk management approach in line with the ISO/IEC standard 31000 assuming that an event of D&FN being investigated is a specialisation of digital risk, and thus it should undertake the same standard approach when dealing with a potential high-stake crime. Ultimately, the component aims at supporting LEAs in their decisions with regard to countering disinformation online as well as their potentially adverse effects on crime itself and the society, doing so by grasping community resilience and suggesting actionable and/ or preventive list of countermeasures.

It is important to note that countermeasures are chosen using an integrated Multicriteria Decision Analysis (MCDA) model, which is implemented through the Analytic Hierarchy Process (AHP). This approach assists decision-makers in evaluating and prioritising multiple competing alternatives based on various criteria. The process involves breaking down the decision problem into a hierarchy of goals, criteria, sub-criteria, and alternatives. Pairwise comparisons are conducted to determine the relative importance of each element at every level of the hierarchy using a numerical scale, and the results are synthesised to calculate weighted scores.

[1] OECD, *Trust in Government* (trust indicators page). Available at: https://www.oecd.org/en/topics/trust-in-government/
[2] European Parliament, *Flash Eurobarometer - Media & News Survey 2022*, Brussels, 2022. Dataset: https://data.europa.eu/data/datasets/s2832_fl011ep_eng
[3] Edelman, *Trust Barometer* (Annual Report). Available at: https://www.edelman.com/trust-barometer

2.1 Function of the Behaviour Profiler & Socioeconomic Analyser for the FERMI Platform

In parallel to the scientific elaboration of the Behaviour Profiler and the Socio-economic Analyser, their communication and explanation of their function to LEAs and related stakeholders took place during the planned FERMI three pilots. Following the feedback collected after the first round, the presentation of the purpose and function of the Behaviour Profiler and the Socio-economic Analyser have been redrafted in order to improve intelligibility. Specifically:

With regard to the Behaviour Profiler following explication text has been drafted for the platform: "The FERMI Community Behavioural Profiler is a measure of community vulnerability towards fake news and disinformation campaigns and is intended to help LEAs and public security policy makers to take appropriate preventive or mitigating measures. It expresses the probability of FIMI operations to find fertile ground in a given societal context, in order to exploit online disinformation for criminal purposes in the offline world. The Behaviour Profiler draws evidence from two key proxy variables: First, Citizens' Media Literacy, and, second, Citizens' Trust in public/private institutions."

With regard to the Socio-economic Analyser, for user requirement 016 ("The user is able to quantify the economic impact by making an approximation on the costs of violent extremism caused by disinformation and fake news") following explication text has been drafted for the platform: "The Socioeconomic Analyzer (SA) operates through a two-stage process, as illustrated in the graph. In the first stage, SA estimates a cost coefficient that captures the impact of increased crime on economic welfare. This is done using actual data on extremist crimes and economic indicators. The econometric model incorporates both the number of extremist crimes and the economic welfare levels across different geographical areas to estimate this coefficient. In the second stage, the estimated coefficient is combined with predicted crime data, linked to disinformation campaigns, to calculate the resulting economic cost in euros for each region."

For user requirement 018 ("The user is able to determine the economic factors that play a role in the ramifications of disinformation"), following explication text has been drafted for the platform: "The cost of hate crime is calculated using the Socio-Economic Analyzer, as shown in UR016. The results are displayed on the UI through a speedometer, which uses a color-coded scale ranging from green (low impact) to red (high impact). The speedometer aggregates the costs from each geographical region within a country to compute a national average. It then assesses whether this average is high in relation to the regional cost distribution. High impact means that the regional average is high compared to the regional cost. This component helps the agency understand how extremist crimes affect the economy by providing a measurable cost. Based on this information, the agency can make more informed decisions about where to allocate resources, for example, by directing more funding or personnel to regions with higher economic impacts from such crimes."

Additionally, a graph with a flow model to visualise the function of the SA has been drafted and featured on the platform prior to the 3rd pilot (see Figure 2: FERMI Socioeconomic Analyser Flow Model.).

The above modifications led to satisfaction of the users during the 3rd pilot which exceeded the set KPIs. More specifically, over 70% of pilot participants agreed that "The citizen is able to increase his/her knowledge about the socioeconomic and cultural aspects and the perception of disinformation among citizens" by using the platform. This is well in line with expectations, considering that this demand was deemed to be of mid-level importance requiring an approval rate of 65%. The above-mentioned user requirements 016 and 018 were clearly met, too, as well over 80% of pilot participants expressed their satisfaction with the model's performance in that regard.

3 Component Description, Implementation and Interaction

The Behaviour Profiler, the & Socioeconomic Analyser and the Community Resilience Management Modeler as all FERMI technical components, are interlocked with each other. The components exchange information and data with each other forming a path starting from analysing and evaluating online content, analysing influencing factors and the impact of disinformation on crimes in the physical, i.e. non-online, world, visualising risks to formulating countermeasures. Each component combines individual data sets, theoretical expertise and technical solutions by the respective partner. The components and the interconnections are depicted in Figure.

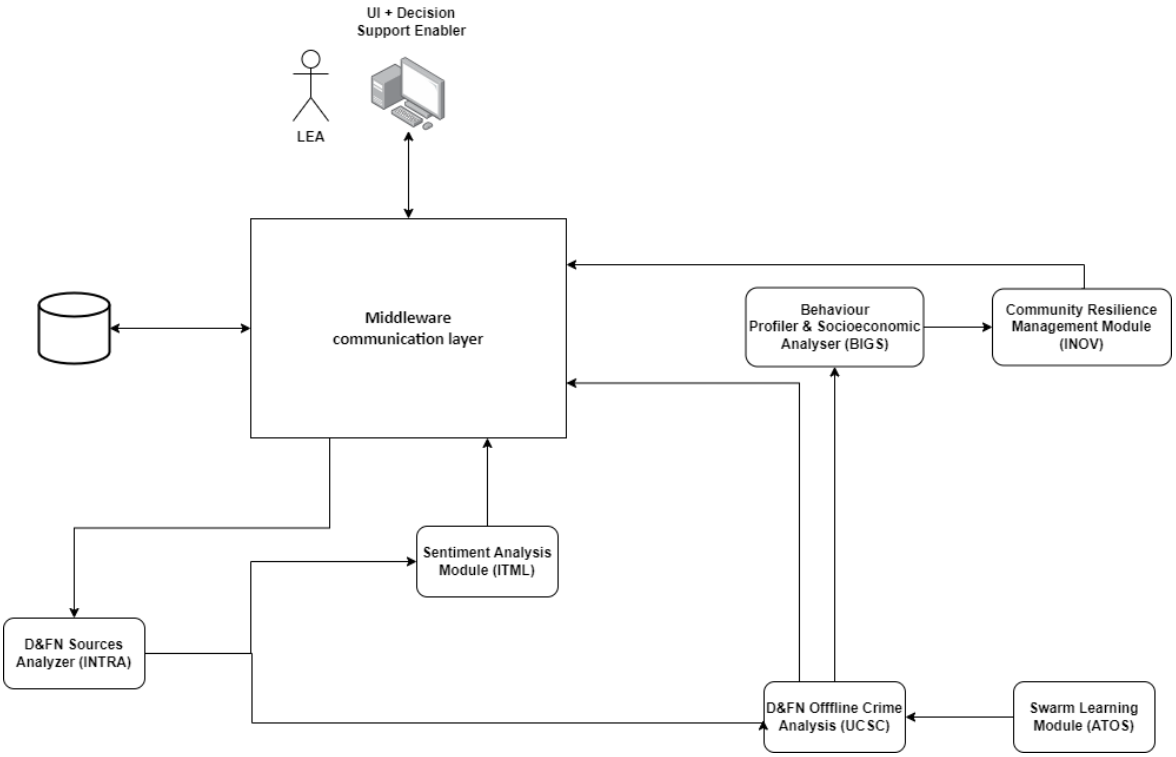

Figure 1: FERMI Technical Components. Source: Deliverable 2.1

Without going into further details about all of the individual components it seems necessary to present where the Behaviour Profiler & Socioeconomic Analyser as well as the Community Resilience Management Modeler receive their input from and where their input goes to. Each box represents a component in the FERMI platform, and each arrow describes a flow of data.

Basically, the process starts at the upper left with the user starting an investigation process through the UI. The Spread Analyser by INTRA uses this input with a specified investigation ID and a tweet or post from which to start fetching corresponding data. With a created dataset from X/Twitter or Mastodon, it identifies bots or human-created content and designs a graph for each sort of investigation that is considered by the FERMI project. This includes investigations related to left-wing, right-wing and COVID-19-related forms of extremism. It discerns disinformation-spreading accounts as well as their influence.

The Sentiment Analysis component further elaborates on this analysis by examining the sentiment of online content that may contain disinformation and thereby enhancing the analysis of online activities and offline criminal behaviour. UCSC's Dynamic Flows Modeler creates a prediction of crimes that occur due to a specific form of disinformation rooted in left-wing, right-wing or COVID-19-related topics in the online content while guaranteeing data privacy it predicts crime caused by disinformation with respect to specific needs of the LEAs. The Swarm Learning Module, moreover, builds a decentralised format that enables work on aggregated rather than local models for collaboration between LEAs in a safe and secure manner.

Finally, the output from the Dynamic Flows Modeler is integrated into the Behaviour Profiler & Socioeconomic Analyser. This is fully in line with the GA's requirement to "integrate information from multiple sources", which includes the analysis of information stemming from social media data and "information about D&FN [disinformation and fake news] sources" in the sense of the social media post's extremist cause's ideological origin.[4] The Behaviour Profiler & Socioeconomic Analyser elaborates on crime that occurred due to disinformation analysing likelihood and severity in greater detail. In addition to technical modelling, the Behaviour Profiler incorporates societal indicators from country profiles, particularly media literacy levels and public trust metrics. Research suggests that societies with lower trust in media and institutions are more vulnerable to disinformation, which can contribute to extremist narratives and, in some cases, lead to disinformation-induced criminal actions. By integrating these societal insights, the Behaviour Profiler enhances its ability to assess not only the likelihood of crime occurring but also how societal vulnerabilities influence the spread and impact of disinformation-fuelled extremism. The integration of this input is further elaborated on in chapter Model Description.

Moreover, what should be taken from the figure is that the input, i.e. the added value, from the Behaviour Profiler & Socioeconomic Analyser goes to the Community Resilience Management Module by INOV. This component uses the inputs to depict likelihood and severity of crimes occurring due to disinformation, checks if the event is of high or low impact and, in case of a high impact event, provides – to quote the GA again – "suggestions for measures that can be taken proportionally to minimize the consequences and the risks."[5]

[4] See 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART B, p. 12

[5] Whilst the GA assigns this requirement to the Behaviour Profiler See ('Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' European Research Executive Agency, 2021, PART B, p. 12), it is fully covered by the platform, albeit by the subsequent Community Resilience Management Modeler.

4 Behavioural Profiler & Socioeconomic Analyser

The FERMI project is premised upon the hypothesis that activities in the online realm can translate into physical real-life actions and impacts. Starting point for the FERMI project and for the Behaviour Profiler & Socioeconomic Analyser is, therefore, to analyse the connection between the online spread of disinformation and various crimes informed by the Dynamic Flows Modeler's results and mediated by media literacy and trust. Moreover, the specificity of the Behaviour Profiler & Socioeconomic Analyser is to narrow down the analysis to a certain form of crimes, namely, to focus on extremist crimes, which are likely to feature particularly prominently in the online disinformation-offline crime nexus, as the overview above has revealed.

4.1 Conceptualisation of the Component

The output of the Dynamic Flows Modeler is supplemented by an assessment of the likelihood that politically motivated crime unfolds and an estimate of the socioeconomic costs thereof, which can further contribute to giving LEAs a better understanding of the consequences of disinformation-induced crime. Accordingly, the outputs of the above-mentioned components are merged into a single solution that includes a measurement of likelihood of politically motivated crime taking place and the severity of occurring politically motivated crimes. The product of these parts gives a measurement of the costs of extremism reflecting the risk connected to disinformation-induced politically motivated crime, which meets the GA-required examination of different "forms of politically motivated extremisms as well as its impact on society […]".[6]

The factor of media literacy may have a chilling impact on the spread of "online propaganda",[7] in other words, extremism-rooted disinformation campaigns, which is not new. Analysing and comparing media literacy and media literacy education in European countries and the US, for instance, has already been under way.[8] Definitions of what media literacy actually entails may be different. An example can be retrieved from Livingstone (2004), where she defines media literacy as the "[…] ability to access, analyse, evaluate and create messages across a variety of contexts […]".[9]

Some studies were able to show that network heterogeneity of social media users is connected to news sharing as well as political interest of the users while political interest and further dissemination of news interact with each other.[10] Thus, political engagement might also be reflected in news sharing activities.[11]

---

[6] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART A, p.9. – Interestingly, the question of identifying the roots of these different forms of extremism ("the causes" and "emergence") also comes up. As explained in the preceding footnote, it is beyond the scope of this project to develop a platform that can detect causes of extremist activities, which are at full display already in the form of a Tweet's ideological message the platform is fed with by the end-user, but the spread analyser can trace social media messages informed by such attitudes back to the original account they emerged from and can further distinguish between human- and bot-originated roots.

[7] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART A, p. 9.

[8] Grafe, S. (2011). ''media literacy' und 'media (literacy) education' in den USA: ein Brückenschlag über den Atlantik.' *Medienpädagogik.* Available at: https://doi.org/10.21240/mpaed/20/2011.09.13.X.

[9] Livingstone, S. (2004). 'Media literacy and the challenge of new information and communication technologies.' *Communication Review*, p.2. Available at: http://eprints.lse.ac.uk/1017.

[10] See Choi et al. (2015). Choi, J. & Lee, J.K. (2015). 'Investigating the effects of news sharing and political interest on social media network heterogeneity.' *Computers in Human Behavior.*

[11] There are approaches in explaining behaviour of people while using social media and sharing news on social media. Karnowski et al. (2018) use the theory of reasoned action that aims at describing behaviour from the origins of that behaviour. The findings suggest (with methodological restrictions and limitations of course) that news sharing in social media is mainly motivated by seeking information and seeking social contact. Moreover, higher news sharing activities is also connected to a higher political interest of social media users. Still, the literature argues that several aspects in explaining behaviour to be considered is – among others – attention, the form of reasoning, perceived source credibility, prior knowledge etc. See Karnowski, V. et al. (2018). 'Why Users Share the News: A Theory of Reasoned Action-Based Study on the Antecedents of News-Sharing Behavior." *Communication Research Reports.* Available at:

Considering media literacy, there is tentative evidence of people receiving media and information literacy training being more likely to be able to estimate the accuracy of information while being less likely to share disinformation.[12] Building on this foundation, the country profiles updated in this deliverable further assess the relationship between media literacy, trust levels, and disinformation resilience across different FERMI relevant European contexts. Beyond adding trust indicators, this version (i) refreshes and harmonises the country profiles and indicator definitions (including media-literacy and press-freedom measures), (ii) adds a comparative synthesis highlighting Key Country Insights, and (iii) documents the integration of country-profile indicators within the FERMI platform (Behaviour Profiler & Socioeconomic Analyser). The inclusion of trust indicators from the OECD, Eurobarometer, and Edelman Trust Barometer provides deeper insights into how societal trust influences the effectiveness of disinformation campaigns. The data suggest that societies with lower trust in media and institutions exhibit higher engagement with alternative or extremist narratives, which can escalate into real-world actions. [13,14, 15] By integrating these societal indicators into the Behaviour Profiler, the FERMI project enhances its ability to potentially predict disinformation-induced politically motivated crimes and assess their broader social impact.

4.2 Model Description

The work of the Behaviour Profiler & Socioeconomic Analyser is described in the GA under task T3.5 and is part of the technical implementation in the FERMI project.[16] In the centre of the model description for both the Behaviour Profiler & Socioeconomic Analyser with its main outcome stands the following equation:

Equation 1: Computation of the Costs of Extremism

$$Crime\_Cost_{c,r,f,t} = Number\_Crimes_{c,r,f,t} * Population_{r,t} * Cost\_Parameter_{c,r,f,t}$$

The model describes the computation of the costs of extremist crimes shown on the left-hand side. These costs are dependent on the following variables: first, the number of crimes following a disinformation campaign and accounting for the nature of crime (c)[17], the region where it is committed (r), the time (t) and a political flag (f) that is connected to an underlying political motivation for a disinformation campaign.[18] The second essential variable of the equation is the cost parameter. This factor is also characterised by nature of crime (c), region (r), time (t) and political flag (f) to guarantee that both factors, i.e., the number of crimes and the cost parameter, are analysed on the same level. This parameter reflects the change of crime costs, measured

https://anna-kuempel.de/publication/karnowski-why-2018/karnowski-et-al_2018.pdf, Obadă, D. & Dabija, D. (2022). '"In Flow"! Why Do Users Share Fake News about Environmentally Friendly Brands on Social Media?' *Environmental Research and Public Health.* Available at: https://doi.org/10.3390/ijerph19084861 and Pennycook, G. & Rand, D (2021). 'The Psychology of Fake News.' *Trends in Cognitive Sciences*. Available at: https://doi.org/10.1016/j.tics.2021.02.007.

[12] See Adjin-Tettey, T. (2022). 'Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education.' *Cogent Arts & Humanities.* Available at: https://doi.org/10.1080/23311983.2022.2037229. More specifically, radicalisation online might be facilitated by characterising in- and out-group in line with radical groups' preferences. The purpose of radical groups online was identified as that their group is being seen in a positive light and simultaneously framing people not belonging to their group in a negative light with further potential of conflict. See Williams, T. et al. (2022). 'How does language influence the radicalisation process? A systematic review of research exploring online extremist communication and discussion.' *Behavioural Sciences of Terrorism and Political Aggression.* Available at: https://doi.org/10.1080/19434472.2022.2104910.

[13] OECD, Trust in Government (trust indicators page). Available at: https://www.oecd.org/en/topics/trust-in-government/

[14] European Parliament (2022) - Flash Eurobarometer: Media & News Survey 2022, European Parliament, Flash Eurobarometer - Media & News Survey 2022, Brussels, 2022.

[15] In a survey conducted in 2022, the Eurobarometer, indices for...ed from and is available at: https://doi.org/10.4232/1.14012.

[16] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021.

[17] Nature of crime is defined here for the crimes that we are looking at, e.g. theft, assault and intimidation.

[18] In the FERMI project, three different categories of disinformation are defined: right-wing, left-wing and COVID-19-related (translated to the crime dataset as "other crime") disinformation campaign.

in economic terms, namely GDP per capita, due to a one unit increase in a crime of specific nature (c), region (r), time (t) and political flag (f).

The cost parameter is computed by the second component, namely the Socioeconomic Analyser. However, it needs to be added that the parameter will be computed as one single coefficient for all NUTS-2 regions. Thus, the parameter itself will be the same for all NUTS-2 regions and all years but will be based on the aggregated results from the individual NUTS-2 regions and years to guarantee data protection. This will be further elaborated on in sub-chapter Socioeconomic Analyser.

Lastly, the measurement of population is the third variable of the equation and it ensures to estimate and output crime costs that have intuitive explanatory power. It is characterised by a specific region (r) and time (t).

It is in equation 1 that the integration between technologies is most relevant. The **Number of Crimes, in this equation, is sourced from the Dynamic Flows Modeler**, which provides a level of offline crime occurrence following a D&FN event online, indicated as **likelihood**. As shown previously in Figure, the image presents the data flow between the FERMI components, as a whole. Within said figure, the linkage between the Dynamic Flows Modeler (D&FN Offline Crime Analyses) and the Behaviour Profiler & Socioeconomic Analyser can be seen. Just as well, the outputs of the Behaviour Profiler & Socioeconomic Analyser are then passed to the Community Resilience Management Modeler.

In the following, the above-described components will be looked at in further detail starting with the Behaviour Profiler with its two sub-tasks and the Socioeconomic Analyser. It offers the logic behind the proposed models and provides the foundation for the technical implementation in chapter Technical Implementation.

While the Behaviour Profiler focuses on crime prediction following disinformation campaigns, societal factors play a crucial role in shaping disinformation susceptibility and the likelihood of real-world consequences. As explained above, the country profiles, which assess levels of media literacy and trust in institutions, provide an essential contextual layer for interpreting the model's predictions. As research suggests that in societies with lower trust in media and institutions, disinformation campaigns are more likely to escalate into offline actions, including politically motivated crimes.[19,20] Thus, integrating these societal indicators could help refine the overall assessment of disinformation-driven extremism and support a more comprehensive predictive approach.

4.2.1 Behaviour Profiler

The Behaviour Profiler consists of two sub-tasks that partially complement each other. The first sub-task is concerned with crime prediction following a disinformation campaign. It is the main component in the technical implementation for the computation of the likelihood of crimes occurring due to a disinformation campaign.

The second task is in respect with the country profiles that have been created. These profiles complement the Behaviour Profiler in a significant way via their inspection of media literacy and information consumption behaviour in the countries of interest for the FERMI project. By analysing trust levels in institutions, media consumption habits, and disinformation resilience, the country profiles provide crucial context for the Behaviour Profiler's predictive capabilities. The inclusion of Spain in this analysis further refines this perspective, as it introduces insights from a country with relatively low institutional trust and high

[19] Marin Lessenski, *How It Started, How It is Going: Media Literacy Index 2022*, Policy Brief 57, October 2022, Open Society Institute – Sofia. Available at: https://osis.bg/?p=4243&lang=en
[20] OECD, *Trust in Government* (trust indicators page), available at: https://www.oecd.org/en/topics/trust-in-government/

political polarisation.[21, 21] By integrating these societal indicators, the Behaviour Profiler can better assess how disinformation campaigns influence public sentiment and, potentially, criminal activities.

4.2.1.1 Crime Estimate

The crime estimation technology, specified in the GA as task 3.1, D&FN-induced and -enabled offline crime analysis and prediction, delivers on its GA commitment to "evaluate the degree in which the spread of D&FN online impacts on the occurrence of offline crime,"[22] and "produce AI-based predictions of the most likely spatiotemporal evolution of D&FN-induced and D&FN enabled offline crimes."[23] Implementing machine learning (ML) methods is an advancement in the standard practices of contemporary policing.

The component developed within FERMI task 3.1, the Dynamic Flows Modeler, was trained to create informed estimates of offline crime occurrences, allowing for the impact of an online disinformation or fake news event in NUTS-2 regions of Europe to be better understood and, in turn, provide insight to LEAs when deciding how to allocate their resources. The output of the model is an estimate for four types of crime's evolution in a four-week period in a specified country's NUTS-2 region that follows the launch of an investigation on the FERMI platform. The Dynamic Flows Modeler's estimates are wholistic, meaning they provide a likely level of crime considering a range of factors, primarily the historical levels of crime (provided via FERMI's Swarm Learning framework) and the online spread of the disinformation post being investigated (delivered by the Spread Analyser). A more in-depth and expansive review of the technology is also provided in Deliverable 3.2, the "technology facilitator package – 2nd version."

The Dynamic Flows Modeler uses AI-driven ML, particularly deep learning, to study a provided time period, with its corresponding data, to understand the patterns and evolution of the provided variables, from which it estimates how one of these variables (crime) will evolve given the provided evolution of all others. Specifically, the DFM is focused on predicting the evolution of offline crime occurrences given the socioeconomic variablesrovided in the datasets and the D&FN's intensity, provided by the platform, in the time period for which it is predicting. While previously the Dynamic Flows Modeler relied on two ML architectures, the 2nd version has been streamlined to employ 1-dimensional convolutional neural network (1D CNN), exclusively.

Whilst a single component, the DFM is comprised of 8 different deep learning models, two for each of the four crime types: one covers the political extremism use-cases, and the other one covers the public health threat use-case. Thus, there is a model specifically trained to predict each unique crime, having studied said crimes occurrences and the level of the spread of disinformation related to the given use cases. Therefore, the models are the following: (1) assault – political extremism; (2) assault - COVID-19/public health; (3) destruction/damage/vandalism of property – political extremism, (4) destruction/damage/vandalism of property – COVID-19/public health; (5) disorderly conduct – political extremism; (6) disorderly conduct - COVID-19/public health; (7) larceny/theft – political extremism, and (8) larceny/theft – COVID-19/public health.

4.2.1.2 Country Profiles

The Country Profiles analysis remains a key component of Task 3.5: FERMI Behaviour Profiler & Socioeconomic Analyser, particularly in understanding how media literacy influences societal resilience to disinformation. The first version, D3.3, focused on, as described in the GA, the ambition "to determine effects of online propaganda on offline actions. In this respect, the degree of media literacy may tend to correspond to the degree of resilience of a society. The means of information and news consumption is a first indicator for

[21] European Parliament, *Flash Eurobarometer - Media & News Survey 2022*, Brussels, 2022. Dataset: https://data.europa.eu/data/datasets/s2832_fl011ep_eng
[22] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021.
[23] Ibid.

the assessment of media literacy. Factors such as the type of source, the "general" assessment of the medium, the level of trust (if feasible) and differentiation by age groups (demographics) play an important role. Based on secondary literature, an analysis of the media literacy of certain countries will be conducted, considering the factors mentioned above. This preliminary work allows behavioural profiles to be better differentiated and classified."

From the first version of the deliverable (D3.3), the analysis of country profiles was structured around media literacy as an indicator of societal resilience to disinformation. The following key aspects were established in that version:

- Media literacy as a resilience factor: (i) In D3.3, it was confirmed that higher levels of media literacy correlate with lower susceptibility to disinformation. (ii) Through respective analysis, it was established that media consumption habits, trust levels, and age-based differentiation play an important role in how societies resist online propaganda.
- Findings on information consumption and trust: (i) It was established that the means of information consumption (TV, online news, social media, etc.) influenced media literacy and trust in news. (ii) Demographic differences were considered where data was available. (ii) Some countries showed clearer vulnerabilities due to low trust in media or high exposure to disinformation.

For this second and final version, we extend and refine the analysis with:

- New indices incorporated: Expanding beyond the previous Press Freedom Index and Media Literacy Index, this version adds trust indicators from the OECD, Edelman Trust Barometer, and updated Eurobarometer surveys.
- The inclusion of Spain: Following the amendment to the GA, which removed the French Ministry of Interior (DMIA) as a project beneficiary and introduced the Spanish Ministry of Interior (GUCI) as a new partner, this version of the deliverable reflects the updated consortium composition by relevantly including Spain in the country profiles analysis and removing France. This adjustment ensures alignment with the project's current structure while maintaining consistency in assessing the effects of media literacy on disinformation resilience across the core FERMI relevant countries.
- Refined media trust and literacy analysis: (i) Press freedom trends (RSF 2024 updates) are integrated to show changes since the previous analysis. (ii) Impact of trust levels on susceptibility to disinformation is further elaborated, particularly with the newest OECD Drivers of Trust survey results and other relevant reports and sources.

In this version, we do not just repeat what was already established in D3.3 but rather update and expand the findings with new data sources, a refined methodology, and an additional country profile relevant for the FERMI context.

4.2.1.2.1 Media Literacy, Trust and Disinformation

This section starts by exploring the crucial role of public trust in societal resilience, highlighting its impact on disinformation trends. As stated in the OECD Trust Survey, "public trust is a pillar of democracy, fostering debate and participation, encouraging compliance with the law, and facilitating reforms. In the face of major environmental, demographic, fiscal and technological challenges, governments need to step up their efforts to strengthen trust, govern effectively and secure democratic resilience […] To do so, governments must ensure that effective public services respond to the needs of a diverse population, policy decisions are

transparent and made in the best interest of the people, checks and balance among institutions are in place, and people can meaningfully participate in decisions." [24]

The latest Eurobarometer reveals the highest level of trust in the European Union since 2007. On a general note, the survey shows that Europeans have a more optimistic view about the future and that they would like to see a stronger and more independent EU, especially in the face of the current global challenges. This is evident from the fact that 51% of Europeans tend to trust the EU, and trust in the EU is highest among the young people aged 15-24 (59%). In another 17 year-record, 51% of Europeans said they trust the European Commission. Almost three quarters of respondents (74%) say they feel as citizens of the EU, the highest level in more than two decades. In addition, more than six in ten EU citizens (61%) are also optimistic about the future of the EU.[25]

The recent OECD (2024) findings illustrate significant variance in public trust levels across countries of interest for FERMI, these are Finland, Sweden, Belgium, Germany, and Spain. Finland and Belgium demonstrate higher trust in both media and governmental institutions, correlating with stronger resilience against disinformation, whereas lower trust levels in Germany and Spain suggest increased susceptibility. The Eurobarometer (2024) further highlights media trust disparities, with Finnish media achieving a 72% trust rate, compared to only 34% in Spain, indicating possible critical areas for targeted interventions.

To better understand how these factors play out in different media landscapes, the following section presents country-specific analyses based on the latest available trust indices, media literacy rankings, and press freedom evaluations.

Key Country Insights:

Finland

- Trust in the national government remains above the OECD average at 47%, though this reflects a considerable decline since the previous survey. [26]
- News media maintains strong credibility, trusted by 61% of Finns[27]
- Exceptionally high trust levels in police (87%) and courts/judicial system (74%), reflecting overall strong institutional trust despite recent governmental declines[28]
- Challenges remain, including pronounced gaps in trust by political agency, age, and gender[29]
- Finland also leads Europe in media literacy (ranked 1st with a score of 74[30]) and has one of the world's freest press environments (5th globally in press freedom[31]). This high-trust, high-literacy environment correlates with stronger resilience against fake news and disinformation

[24] https://www.oecd.org/en/topics/trust-in-government.html
[25] https://ec.europa.eu/commission/presscorner/detail/en/ip_24_6126
[26] https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_a8004759-en/finland_596ba5da-en.html
[27] ibid
[28] ibid
[29] ibid
[30] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf
[31] RSF – Reporters without borders, *Finland*. Available at: https://rsf.org/en/country/finland

Sweden

- Moderate to high trust in the national government is at 43%, marking a positive trend from previous years[32].
- Media trust is relatively stable (45%), indicating moderate public confidence[33].
- Swedish citizens exhibit high trust in core societal institutions like police (69%) and the judiciary (64%)[34].
- Political disenfranchisement notably influences trust, creating a substantial 49% point gap[35].
- Sweden ranks among the top in media literacy (5th in Europe, score 71[36]) and maintains a very free press, Sweden improved to 3rd place in 2024, up from 4th in 2023 in press freedom[37]. The Edelman Trust Barometer finds Sweden's overall trust index (average percent trust in NGOs, business, government, and media) at 49 (out of 100), The relatively robust institutional trust and literacy in Sweden can help inoculate the public against misinformation.

Belgium

- Trust in the federal government is notably high at 47%, significantly above the OECD average.
- News media trust is comparatively strong, at 52%, highlighting good public confidence in media reliability.
- Institutions like the police (67%) and the judicial system (53%) are significantly more trusted compared to political parties (29%).
- However, considerable gaps based on financial concerns and educational attainment reveal vulnerabilities that could influence susceptibility to disinformation.
- Belgium's media literacy levels are strong (ranked 10th in Europe, score 61[38]). Belgium made a significant jump to 16th in 2024 in the press freedom index (from a relatively low 31st in 2023) recovering from the previous year's decline[39]. Overall, Belgium's rising trust and solid media-literacy foundation suggest improved resilience to disinformation.

Germany

- Trust in the federal government is below the OECD average, at 36%[40].

[32] https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_a8004759-en/sweden_11ca1946-en.html
[33] ibid
[34] ibid
[35] ibid
[36] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf
[37] RSF - Reporters without borders, *Press freedom index - 2024*. Available at: https://rsf.org/en/country/sweden
[38] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf
[39] Reporters Without Borders, *Word Press Freedom 2024 Index: Belgium Country Profile*, https://rsf.org/en/country/belgium
[40] https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_a8004759-en/germany_9f727e4a-en.html

- The German public shows notably low trust in the news media (34%), suggesting higher susceptibility to disinformation.[41]
- Conversely, trust is notably higher in law enforcement (police: 64%) and the judicial system (58%)[42].
- A major concern is the political disenfranchisement gap, with trust levels 54 % points lower among those feeling politically voiceless[43].
- The country's media literacy score (61, 11th in Europe) is on par with Belgium's, and its press remains fairly free, Germany similarly climbed to 10th in 2024 after dropping to 21st in 2023 in the press freedom index, restoring it to the top 10[44]. Yet, the Edelman Trust Barometer finds Germany's overall trust index (average percent trust in NGOs, business, government, and media) at just 45 (out of 100), which can be indicative of general public scepticism[45] This trust deficit, despite good education and press freedom, suggests that significant segments of the German public may be more susceptible to disinformation narratives that exploit institutional distrust.
-

Spain

- Trust in the central government stands at 37%, slightly below OECD average[46].
- Trust in the news media is similarly low (34%), indicating potential challenges in countering disinformation[47].
- Public confidence remains stronger in institutions such as the police (61%) and judiciary (45%)[48].
- Trust gaps across education levels are significant, with lower trust levels among less educated groups, potentially heightening vulnerability to disinformation.
- Spain's media literacy ranking is 16th in Europe (score 58), trailing its Northern European peers, and its press freedom stands at 36th worldwide after a recent decline. Spain remains among the 30% of countries that Reporters Without Borders (RSF) deems to have a "fairly good" situation for press freedom ranked 30th, dropping from 36th in 2023[49]. The Edelman Trust Barometer finds Spain's overall trust index at 46 (out of 100)[50]. This combination of lower trust and literacy can signal a higher vulnerability to disinformation.

Cross-Country Analysis & Implications

The analysis of trust in institutions, media literacy levels, and press freedom across Finland, Sweden, Belgium, Germany, and Spain provides key insights into societal resilience against disinformation. The data highlights differences in media ecosystems and public trust, which influence how each country experiences and mitigates disinformation risks.

---

[41] ibid
[42] ibid
[43] ibid
[44] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva*, "Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf
[45] https://www.edelman.com/sites/g/files/aatuss191/files/2024-02/2024%20Edelman%20Trust%20Barometer%20Global%20Report_FINAL.pdf
[46] https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_a8004759-en/spain_56998449-en.html
[47] ibid
[48] ibid
[49] Reporters Without Borders, *Word Press Freedom 2024 Index: Spain Country Profile*, https://rsf.org/en/country/spain
[50] https://www.edelman.com/sites/g/files/aatuss191/files/2024-02/2024%20Edelman%20Trust%20Barometer%20Global%20Report_FINAL.pdf

- **High-Trust, High-Literacy Countries: Finland and Sweden**
    - Finland and Sweden exhibit the highest resilience to disinformation due to a combination of strong institutional trust, high media literacy scores, and free press environments.
    - Finland leads Europe in media literacy (1st place, score 74) and enjoys one of the world's freest press environments (5th globally in 2024). While trust in the national government (47%) has slightly declined, trust in news media remains high (61%), reinforcing a stable information environment resistant to manipulation.
    - Sweden ranks 5th in media literacy (score 71) and has improved its press freedom ranking to 3rd place in 2024. Trust in the government (43%) and media (45%) remains moderate to high, and Sweden's relatively high levels of public trust and strong educational foundation contribute to its ability to counter disinformation effectively.

  Implications
    - Both countries demonstrate that a high-trust, high-literacy environment strongly correlates with societal resilience against disinformation.
    - Policy measures reinforcing transparency, civic engagement, and digital education are effective strategies to maintain this resilience.

- **Moderate-Resilience Country: Belgium**
    - Belgium falls into a moderate resilience category, benefitting from relatively strong media literacy (ranked 10th score 61) and a recent improvement in press freedom (16th in 2024, up from 31st in 2023).
    - Government trust (47%) and news media trust (52%) are moderate to strong, indicating relatively high public confidence in institutions.
    - However, financial concerns and educational disparities create vulnerabilities that could influence susceptibility to disinformation.

  Implications
    - While Belgium's overall trust and literacy levels support a resilient society, economic and educational divides could be beneficial to be addressed to ensure uniform disinformation resistance across demographics.
    - Targeted digital literacy programs in lower-income and lower-education groups could further strengthen Belgium's disinformation resilience.

- **Lower-Trust, Moderate-Literacy Country: Germany**
    - Germany presents a contrasting case, despite a strong education system and a relatively free press (ranked 10th in 2024, up from 21st in 2023), public trust levels are low. Trust in government (36%) and media (34%) is one of the lowest among Western European democracies, indicating widespread scepticism toward institutions.
    - Media literacy is moderate (ranked 11th, score 61), yet a significant political disenfranchisement gap (54 percentage points) highlights concerns over democratic engagement.

  Implications
    - Germany's trust deficit, despite strong press freedom and literacy levels, suggests an increased vulnerability to disinformation that exploits institutional distrust.

- Rebuilding public trust through transparency initiatives and reinforcing digital literacy could be found critical in addressing this susceptibility.

- **Low-Trust, Lower-Resilience Country – Spain**
  - Spain ranks lower in resilience compared to Northern and Western European peers. Its low government (37%) and media trust (34%) levels combined with moderate media literacy (ranked 16th, score 58) indicate higher susceptibility to misinformation.
  - Press freedom (30th in 2024, up from 36th in 2023) remains weaker compared to other EU democracies, and political polarisation further exacerbates trust issues.
  - Trust disparities based on education levels highlight the need for targeted interventions to reduce disinformation exposure among vulnerable groups.

  Implications
  - Spain's combination of low trust, moderate literacy, and political polarisation makes it more vulnerable to disinformation, particularly around elections and regional debates.
  - Improving transparency in news reporting, investing in media education, and countering hyper-partisan narratives can be some of the crucial steps to strengthen Spain's resilience.

These findings reaffirm the critical role of trust in institutions and media literacy in shaping societal resilience to disinformation. While the patterns identified align with previous analyses, this updated assessment reflects relevant project (consortium) changes and integrates recent data on institutional trust and media consumption trends, when available. A more detailed examination of these results, including their implications for disinformation resilience, is presented in Section 4.4.1.2.

4.2.2 Socioeconomic Analyser

The Socioeconomic Analyser is mainly covered by task T3.5 and is considered with the connection between crime and effects on economically measurable variables, such as GDP per inhabitant. By applying econometric methods, the effects of radicalization and extremism will be reflected in financial terms to quantify the costs of these negative effects (based on data availability in the respective country/region) for the society.

The following model depicts in a succinct way the inputs, components, function and outputs of the FERMI Socioeconomic Analyser model:

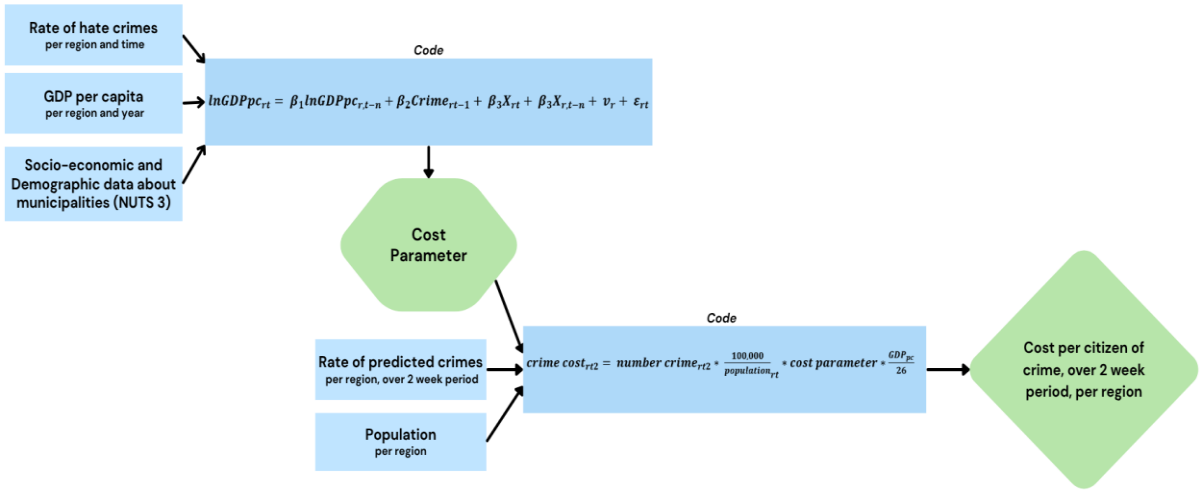

Figure 2: FERMI Socioeconomic Analyser Flow Model. Source: BIGS

The intuition behind this analysis is coming from Ferguson et al. (2019) who built the theoretical foundation for looking at possible effects of political extremism on economic variables. It is proposed that political extremism leads to a loss in social welfare via different channels. These may concern the deterrence of investors, the influence on political decisions and institutional output as well as trade and further factors.[51]

Starting point for the Socioeconomic Analyser is therefore how to approach the topic of political extremism. In this regard, extremism as a concept explaining certain phenomena is hard to define. There is no universal clear-cut explanation of what extremism is, how it is visible and how it is to be detected. Ferguson et al. (2019) took three different approaches to define extremism for the case of Germany. The first concerns a removal or impairment of fundamental values and norms that contradict the German constitution. The second approach views extremism as a stand-alone phenomenon with individual characteristics, thereby focussing more on the counting and definition of the scope of extremism. A third definition describes extremism as a non-acceptable distance of political views from a political centre.[52]

In the Socioeconomic Analyser, the focus is on extremism in terms of extremist crimes. The Analyser uses data on extremist crimes to calculate the economic cost of extremism. To this end, Germany serves as the foundation for the other countries, since it provides data on politically motivated crime, distinguishing between the motivation and the type of crime. Furthermore, these data are also disaggregated at the geographical level.

The definition of extremist crimes is provided by the *Bundeskriminalamt* and the *Verfassungsschutz* in Germany. The latter, representing Germany's domestic intelligence agency, defines right-wing extremism as a movement that is extremist and desires a society where there is clear hierarchy, based on the endowed supremacy of certain individuals or the supremacy of a specific sect of society. Moreover, right-wing extremism aims to achieve its goals by non-democratic means.[53] Crimes defined as right-wing underlie the assumption of inequality of people. If there are connections to nationalism, racism, social Darwinism or National Socialism corresponding crimes are considered right-wing.[54] Left-wing extremism is defined by the Verfassungsschutz as extremism motivated by anarchist or communist ideologies that aim to overthrow the

[51] Ferguson, N. et al. (2019). 'Die Kosten des Extremismus'. *BIGS Standpunkt zivile Sicherheit*. Available at: https://www.bigs-potsdam.org/app/uploads/2020/06/BIGS-Standpunkt_Nr.-9-2019_Kosten-des-Extremismus_WEB.pdf.
[52] Ibid.
[53] Bundesamt für Verfassungsschutz. 'Begriff und Erscheinungsformen'. Available at: https://www.verfassungsschutz.de/DE/themen/rechtsextremismus/begriff-und-erscheinungsformen/begriff-und-erscheinungsformen_artikel.html.
[54] BKA. ‚Politisch Motivierte Kriminalität (PMK) - rechts –'. Available at: https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKrechts/PMKrechts_node.html.

existing state or social structures. Left-wing crime is also based on the assumption of equality of all people, and offenses are often committed due to a lack of belief in the effectiveness of political reforms. Subsequently, to overthrow the capitalist system of a democratic state and the market-based property system, left-wing crimes are committed to enact a political agenda[55]. According to the Bundeskriminalamt, property damage, theft, and violations of the Assembly Act are dominant within left-wing crime[56].

Politically motivated crimes or political extremism that, according to criminal assessment, do not directly fall into the categories of left, right, foreign, or religious ideology. Nevertheless, these criminal acts aim to hinder political objectives or disrupt the free democratic basic order, are generally defined as politically motivated crimes[57]. These definitions are addressed within the FERMI case studies.

Apart from Germany, no other country in the FERMI consortium publicly reports systematic data on politically motivated crimes. However, Spain publishes data on hate crime. According to the official definition, hate crimes are defined as "as any criminal offense, including those committed against persons or property, where the protected legal property is targeted because of its actual or perceived connection, sympathy, affiliation, support or membership of a group. A group is based on a common characteristic of its members, such as their actual or perceived race, national or ethnic origin, language, colour, religion, age, disability, sexual orientation, or other similar factor."[58] These data are reported from 2013 to 2022 for all Spanish regions and provinces.

Hate crimes are conceptually different from politically motivated crimes. However, they can convey the idea expressed by our theoretical approach to extremism and, therefore, Spain serves as the second testbed case, besides Germany, for the exact cost measurement.

The model to explain the economic cost of extremist crimes is depicted in Equation 2. The primary focus of this equation is the calculation of the cost parameter, which indicates how a one-unit increase in extremist crime impacts economic welfare.

Equation 2: The calculation of economic costs to political extremism

$$Prod_{r,t} = \alpha + \beta_1 Ext_{c,f,r,t} + X_{r,t} + v_r + \varepsilon_{r,t}$$

Where $Prod_{r,t}$ describes the measurement of productivity, α is the constant for the regression, v_r is a vector of time-invariant region-specific properties and ε_{rt} is the error term for the regression. $Ext_{c,f,r,t}$ is the measurement of extremism in terms of crime, β_1 gives the cost coefficient, and X_{rt} describes a vector of control variables.

The model was developed to be used for all the countries in the FERMI consortium. However, since Germany and Spain represent two different use cases with different types of data available, we adapt the equation to each country. Thus, in what follows, we present the model first for Germany and then for Spain.

4.2.2.1 Economic cost calculation for Germany: model and data

The main model to explain economic costs by politically motivated crime is described in the following regression equation, recalling

Equation 2:

[55] Bundesamt für Verfassungsschutz. 'Begriff und Erscheinungsformen'. Available at: Federal Office for the Protection of the Constitution - Left-wing extremism
[56] BKA - Politically motivated crime - left -
[57] BMI - Politically motivated crime - Politically motivated crime
[58] The definition can be found at Instructions and Methodology available at https://oficinanacional-delitosdeodio.ses.mir.es/publico/ONDOD/en/publicaciones.html

Equation 3: The economic cost of extremism in Germany

$$Prod = \alpha + \beta_1 Ext_{c,f,r,t} + \beta_2 X_{r,t} + v_r + \varepsilon_{r,t}$$

$Prod_{r,t}$ describes the measurement of productivity for a given region (r) and time period (t), proxied by GDP per capita in thousands of euros. The choice to use this proxy to measure the economic cost is a catch-all approach that reflects the GA's emphasis on numerous socioeconomic factors, as it captures the economic outcome of all such changes. This approach is further corroborated by the GA's reference to the "means of production" and "a loss of productivity" in this regard.[59] α is the constant for the regression, v_r is a vector of time-invariant region-specific properties. ε_{rt} is the error term for the regression. $Ext_{c,f,r,t}$ is the number of politically motivated crimes per ten thousand inhabitants, used as a measure of extremism in terms of crime. This measurement varies by region (r), time (t), political flag (f), and nature of the crime (c). β_1 is the cost parameter indicating how a one-unit increase in extremist crime affects GDP per capita i.e. economic welfare. Lastly, X_{rt} describes a vector of control variables.

Our model controls for time-invariant characteristics for the region under inspection, i.e. this term captures variance that is specific to a given region. This is fully in line with FERMI's ambition to deliver continuous analyses of all the socioeconomic factors that may be location-specific. Regional differences are not uncommon in this field of research. Cultural, historical and social developments might explain different activities and might also influence the perception of people. One example is shown for the case of Germany, where a recent survey found differences in observing extremist violent activities in their own surrounding between the *Neue Bundesländer* and the *Alte Bundesländer* (eastern and western parts of Germany). Survey participants from the former indicated to observe more left-wing and right-wing extremist activities compared to participants from the latter. Furthermore, survey participants from the *Neue Bundesländer* indicated to feel significantly more often threatened by extremist violence.[60]

Still, when analysing economic outcomes, other potential time-variant explanatory factors need to be considered to avoid omitted variable bias, i.e. the risk that the relevant explanatory factors left out in the regression model may potentially distort the findings.[61]

---

[59] The GA alludes to "different areas of economic measurement of the costs of extremism, which can lead, for example, to welfare losses through the loss of human life as well as through the physical destruction or damage of (intellectual) property, assets or other **means of production**" (emphasis added) as well as "costs […] borne by the direct or indirect victims of extremism [that] can affect investment and real estate - as they can deter investors - directly affect policy decisions and institutional outputs, and even affect trade, mobility and tourism. This can lead to job losses - through relocation or expansion to other locations - and consequently a **loss of productivity** [emphasis added] and wealth in a region." See Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART A, p. 9.

[60] Endtricht, R. & Kleinschnittger, J. (2023). 'Wahrnehmung extremistischer Aktivitäten und subjektives Bedrohungserleben in Deutschland: Verbreitung und Einflussfaktoren.' In: *Kriminalität und Kriminologie im Zeitalter der Digitalisierung*, by L. Deyerling, A. Dreißigacker, I. Henningsmeier, M. Neumann, J. Schemmel, C. P. Schröder & L. Treskow T. Bliesener.

[61] Since *certainty* for the researcher with regard to knowing which factors meet the conditions of control variables is highly unlikely careful selection of control variables is needed. Simply put, variables "[…] measured before the variable of interest was determined are generally good controls." (Angrist et al. 2008, p. 50). For the case of the Socioeconomic Analyser, the "variable of interest" is here represented by the measure of economic outcome and control variables should ideally not have an effect on politically motivated crime. See Angrist, J.D. & Pischke, J.-S. (2008). 'Mostly harmless Econometrics: An Empiricist's Companion'. Available at: http://diglib.globalcollege.edu.et:8080/xmlui/bitstream/handle/123456789/141/Angrist%20J.D.%2C%20Pischke%20J.-S.%20Mostly%20Harmless%20Econometrics%20%28PUP%2C%202008%29%28ISBN%20069112034X%29%28O%29%28290s%29_GL_.pdf?sequence=1&isAllowed=y. It has to be added that the inclusion of additional control variables is not always beneficial in respect to what is being researched. Adding unnecessary control variables to the estimation might also increase bias. For a discussion on this topic see for example Angrist, J.D. & Pischke, J.-S. (2008). 'Mostly harmless Econometrics: An Empiricist's Companion'. Available at: http://diglib.globalcollege.edu.et:8080/xmlui/bitstream/handle/123456789/141/Angrist%20J.D.%2C%20Pischke%20J.-S.%20Mostly%20Harmless%20Econometrics%20%28PUP%2C%202008%29%28ISBN%20069112034X%29%28O%29%28290s%29_GL_.pdf?sequence=1&isAllowed=y, Clarke, K. A. (2005). 'The Phantom Menace: Omitted Variable Bias in Econometric Research'. *Conflict management and peace science*. Available at:

Both the components in the Behaviour Profiler & Socioeconomic Analyser use data that is provided by partner institutions (mainly LEAs) and data that can be publicly extracted from websites. Thus, data for Germany are mainly collected from INKAR[62] and EUROSTAT[63]. For the case of crime data, however, some critical remarks need to be stated. Data on politically motivated crime are contained in the dataset on Politically Motivated Crime (Politisch Motivierte Kriminalität – PMK), which is systematically different from the Police Crime Statistics (Polizeiliche Kriminalstatistik – PKS). [64] One difference of particular relevance to the FERMI project is that the PMK dataset classifies crimes according to the political motivation of the perpetrator. As indicated above, categories include right-wing, left-wing, religious, foreign ideology, or not attributable to a specific motivation. [65]

The State Criminal Police Offices (Landeskriminalämter) of the German federal states are responsible for reporting PMK data to the Federal Criminal Police Office (Bundeskriminalamt). However, police offices below the federal state level are not required to publish these data, and in many cases, the details of the PMK dataset are classified. In Germany, data on politically motivated crime are considered sensitive and are therefore subject to strict conditions regarding access and use. As a result, accessing and obtaining publication rights for PMK data below the NUTS-1 level is particularly challenging.

Thanks to its network, BIGS has obtained access to data at the NUTS-3 level. However, it must be ensured that the published results do not reveal sensitive information. Specifically, it is essential to prevent any possibility of reconstructing the actual number of politically motivated crimes in individual NUTS-3 regions in Germany. To address this, a single coefficient is calculated to represent the overall change in GDP per capita resulting from a one-unit increase in politically motivated crime. This coefficient is derived using data aggregated across all NUTS-3 regions, ensuring that no information about actual crime numbers in any specific region is disclosed.

An overview of the data used for the Socioeconomic Analyser can be found in Table 1.

Input (Explanatory Factors)	Controls (Additional Factors)	Outputs (Explained Factors)
Data on politically motivated crime	Fertility rate	GDP per capita
	Employment rate	
	Investment in mining and industry	
	Drop-out rate from school	
	Student leaving school without diploma	

http://www.saramitchell.org/clarke05.pdf, Rosenbaum, P. R. (1999). 'Choice as an Alternative to Control in Observational Studies'. *Statistical Science.* Available at: https://projecteuclid.org/journals/statistical-science/volume-14/issue-3/Choice-as-an-Alternative-to-Control-in-Observational-Studies/10.1214/ss/1009212410.pdf.

[62] INKAR ('Indikatoren und Karten zur Raum- und Stadtentwicklung') comprises socioeconomic data for about 600 indicators on Germany and Europe. For Germany, a more granular dataset is available covering data on NUTS-3 level. Therefore, INKAR provides helpful data sources for the purpose of the Socioeconomic Analyser. Data is available at: https://www.inkar.de/.

[63] Data is available on a NUTS-3 level at: https://ec.europa.eu/eurostat/web/regions/data/database-tbd.

[64] For instance, the *PMK* statistic is an entry statistic, i.e. crimes are reported in this data at the onset of the investigation. On the other hand, the *PKS* statistic is an output statistic, i.e. these crimes will be recorded after the investigation process has ended and files have been handed over to the jury or prosecution. More information on German crimes statistics are available at: https://www.bmi.bund.de/SharedDocs/faqs/DE/themen/sicherheit/pks/pks-und-pmk.html.

[65] BKA. (2024). 'Politisch Motivierte Kriminalität'. Available at: https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/pmk_node.htmlhttps://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/pmk_node.html

Table 1: Input, output and socioeconomic controls utilised for the Socioeconomic Analyser

The input factor is the number of politically motivated crimes per 100,000 inhabitants in NUTS-3 German regions over the period 2010-2021. The types of crimes analysed are theft, disorderly, destruction, assault. These crimes are classified according to four political motivations that is left, right, foreign and religious. The output is GDP per capita of each NUTS-3 region in year *t* in thousands of euros from INKAR. Control variables comprise demographic and economic characteristics of the country. See Table 2 for the list of all control variables.

For the result of the Socioeconomic Analyser, the coefficient for the cost of crime, i.e. the change in economic output (GDP) following an increase of politically motivated crime by one unit is incorporated into a model encompassing the Dynamic Flows Modeler's crime estimates. In the case of Germany, these crime estimates can be assigned to crimes in the German dataset directly according to crime labels. This approach cannot be replicated to other countries participating and represented in the FERMI consortium due to data limitation.

For the case of media literacy that was intended to be included in the Socioeconomic Analyser, restricted data availability was prohibiting the inclusion as an explanatory or control variable in the estimation models. Data on media literacy is available as an index and is created from a fusion of measurements for – among others – scores for the freedom of the press and PISA scores as it has been mentioned in sub-section Behaviour Profiler.[66] Nevertheless, this data is only available on a country level and thereby is not suitable for the inclusion in the regression analysis since all variables need to be measured on the same geographical level to deliver meaningful results.

Proxies for the index of media literacy were created by the EU in form of a measurement for media use and media trust extracted from an EU-wide survey.[67] [68] These measurements give an overview on how and what media is consumed, and it gives an indication for the perception of trust by people with regards to media. However, this data is also incompatible with the regression analysis due to its measurement at the NUTS-1 level and therefore cannot be included for meaningful regression output, either.

As previously mentioned, these coefficients are integrated into the Behaviour Profiler through Equation 1. However, the equation must be adapted to align with the data and predictions used. First, the frequency of offences must be accounted for. While UCSC predicts the absolute number of crimes, the actual number of crimes used in equation (1) is expressed as a frequency. To adjust for this difference and the variation in population sizes across NUTS regions, we scale the predicted number of crimes by dividing it by the region's population and multiplying it by 100,000. Additionally, since the estimated coefficients are based on GDP per capita measured in thousands of euros (as provided in the INKAR database), the cost parameter must also be multiplied by 1,000.

[66] Among these measurements is the Freedom of the Press score, the Trust in Others by the World Values Survey and PISA scores published by the OECD and the World bank. See Lessenski, M. (2023). '*Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*.' Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf

[67] In a survey conducted in 2022, the Eurobarometer, indices for media use and for media trust were created. Data can be extracted from and is available at: https://doi.org/10.4232/1.14012.

[68] The "[…]analysis of the moderation and recommendation systems employed by social media[…]" the GA alludes to ('Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART B, p.12) is covered mainly by the overall analysis of the media landscape by examining media and news consumption behaviour of users of social media. The in-depth and cross-country analysis of the – for FERMI interesting – part of D&FN that considers the illegal realm as well, however, has seemingly become obsolete, considering that these proceedings have been largely standardised by now. In cases of illegal goods, content or services in online spreading, the Digital Services Act (DSA) calls for these platforms to put in place countermeasures, such as "[…]to flag such content and for platforms to cooperate with "trusted flaggers"[…]". See EC. (2023). 'The impact of the Digital Services Act on digital platforms'. Available at: https://digital-strategy.ec.europa.eu/en/policies/dsa-impact-platforms.

Second, the time window must be adjusted. The cost coefficients estimate annual impacts on GDP, whereas UCSC predicts crimes over two-week periods. To align these timeframes, the coefficients are divided by 26, as a year comprises approximately 52 weeks and the two-week prediction period represents half of this frequency. This scaling ensures the coefficients match the temporal resolution of the predictions. Therefore, Equation 1 can be rewritten in this form:

Equation 4: Computation of the Costs of Extremism for Germany

$$crime\ cost_{rt} = number\ crime_{rt} * \left(\frac{100{,}000}{population_{rt}}\right) * cost\ parameter_t * \frac{1{,}000}{26}$$

4.2.2.2 Economic cost calculation for Spain: model and data

The main model to explain economic costs by hate crime in Spain is depicted in the following regression equation recalling

Equation 2 but adapting it to:

Equation 5: The economic cost of extremism in Spain

$$lnGDPpc_{rt} = \beta_1 Crime_{rt-1} + \beta_2 X_{rt} + \varepsilon_{rt}$$

Here $lnGDPpc_{rt}$ is the logarithm of GDP per capita in region r and year t, serving as a measure of economic welfare in Spanish regions. Differently from the German use case in equation 5 the log of GDP is used due to the skewed distribution of GDP. $Crime_{rt-1}$ represents the number of hate crimes per ten thousand inhabitants in region r and year t–1, and is used as a proxy for extremism in the form of crime. This measure varies according to the underlying motivation behind the crimes. In the Spanish case, we estimate the effect of hate crimes in the previous year on economic welfare in the current year, assuming that such crimes take time to impact the economy. Differently from the German case, past crime affects current economic welfare in Spain, as Spanish data are reported differently from the German ones and a lag in their effect is assumed. β_1 is the cost parameter indicating how much a one-unit increase in extremist crime affects GDP per capita. Lastly, X_{rt} describes a vector of control variables and ε_{rt} is the error term for the regression.

Data on hate crimes are publicly available [69] and include the number of hate crimes in each Spanish region (NUTS-2 level) and province (NUTS-3 level), classified by both motivation and type of crime, over the period 2014–2023. Eleven motivations for committing a hate crime are identified: anti-gitanism, anti-semitism, apiphobia, religious beliefs or practices, hate crimes against people with disabilities, generational discrimination, discrimination on the basis of disease, sex/gender discrimination, ideology, sexual orientation and gender identity, and racism/xenophobia. A total of 157 types of crimes that can be committed under these motivations are listed. We use data on hate crimes categorised by motivation and disaggregated at the NUTS-3 level for the period 2014–2023, and we calculate the number of crimes per 100,000 inhabitants.

Data on GDP per inhabitant are sourced from Eurostat and are measured in euros at current market prices. Using information on the Consumer Price Index, also from Eurostat, we convert GDP figures into 2022 constant Euros. Control variables are also obtained from Eurostat and include the employment rate, the percentage of early leavers from education, net migration, the fertility rate, and the number of economically active people.

Input (Explanatory Factors)	Controls (Additional Factors)	Outputs (Explained Factors)

[69] https://estadisticasdecriminalidad.ses.mir.es/publico/portalestadistico/en/datos.html?type=pcaxis&path=/Datos6/&file=pcaxis

Data on hate crime	Employment rate	GDP per capita
	Percentage of early leavers from education	
	Net migration	
	Fertility rate	
	Number of economically active people	

Table 2: Input, Output and socioeconomic factors for Spanish model

Results for Spain are not combined with data from the Behavioural Profiler, as no suitable data were available to predict the number of offline crimes caused by disinformation campaigns.

4.3 Technical Implementation

Along the lines of the Model Description chapter, the technical implementation description will also be divided into the Behaviour Profiler component and the Socioeconomic Analyser. In the upcoming sections the outlined theoretical and descriptive models from the previous sections are translated into a technical instrument.

4.3.1 Behaviour Profiler

With respect to the Behaviour Profiler two technical implementations are provided. In the following, the technical implementation of both will be presented. For the crime forecasting, UCSC is covering the technical implementation qua the ML-driven Dynamic Flows Modeler. Secondly, the technical implementation of the Country Profiles is considered. This part will cover briefly the rather simple implementation via the establishment of a separate website to be accessed by the end-user.

4.3.1.1 Crime Forecast

The methodology undertaken to construct the Dynamic Flows Modeler pursues two primary tasks: (1) pre-processing the data used for training and (2) the development of the ML architecture that produces the forecasts, via studying the behaviour of past data. In this subsection Crime Forecast, the focus is on the latter, that is, the ML architecture. Specifics regarding data pre-processing can be found in Deliverable 3.1 and 3.2, the technology facilitators packages. Said ML architecture was originally comprised of three different approaches and two architectures, on a crime type basis. In the 2nd version, this was re-evaluated, and the component was streamlined to adopt a single architectural approach, 1D-CNN.

1D CNN can be characterised as a deep learning model tailored for processing sequential information that captures relevant patterns and features within data for various tasks. It has proven effective for predictions based on time-series data and typically consists of two fundamental components: the CNN, which extracts and filters the relevant features, and the fully connected layer, which generates the predictions using the features extracted by the CNN, taking for granted their assigned relevancy.

The CNN design used in the Dynamic Flows Modeler includes 3 convolutional blocks, each consisting of varying amounts of filters. In the first version, the initial set employed 500 filters, followed by 250 filters in the second set, and 128 filters in the third set, now these filter sets have been modified to now contain 32, 64, and 128 filters, respectively. The filters apply convolutional operations, enhancing the network's capacity to recognise significant patterns in the data. Rectified linear unit activation was employed in each convolutional layer and the model was trained to minimise the mean squared error loss and a drop-out rate of 0.2 (or, 20%)

was also introduced in order to prevent over-fitting.[70] The model's aim to minimise the mean squared error means it continually adjusts the parameters to improve accuracy and lower the discrepancies between predictions and target value. The following observations present how mean square error was calculated, with n being the total number of data points, y_i the target value for the $i - th$ data point (the real, unseen by the model, observation), and $f(x_i)$ the predicted value produced by the model, for the $i - th$ data point.

4.3.1.2 Country Profiles

The technical implementation of the country profiles has been covered by establishing a separate website within the FERMI platform. This website can be accessed by the end-user and will deliver an overview of additional information on media literacy and news consumption at a national level (which, again, is what data availability allows for) for the countries under consideration. Relevant updated information, where available, will feed the platform and can be found in the respective sections herein.

4.3.2 Socioeconomic Analyser

The technical implementation of the Socioeconomic Analyser is mainly achieved by estimating the model in

Equation 2. This estimation comprised statistical means to comply with the GA which aims at "[…] applying econometric methods […]"[71] to achieve the computation of the effect of crime on economic welfare. For that purpose, the statistical software STATA version 18 was used.

To approach this complex matter, the standard Arellano-Bond Generalized Method of Moments (GMM) estimation method for panel data analysis is applied. It takes into consideration possible endogeneity issues in estimation as well as providing opportunities for including explanatory instrumental variables. Endogeneity in

Equation 2 may arise since crime might possibly predict the income of an economy (GDP per capita). At the same time, however, this might also work vice versa, i.e. that GDP per capita might explain the occurrence of crime. A second concern is that previous values of GDP per capita might correlate with present values of GDP per capita, giving rise to issues with autocorrelation.

The GMM-Arellano-Bond estimator takes these issues into consideration. First, the estimator is applied to datasets with a relatively small number of time periods while capturing a comparably larger number of individuals (e.g., regions, countries etc.). This estimator also takes the issue of endogeneity into account. Moreover, it allows explanatory factors that might be related to past and current values of the error term (and would lead to potential endogeneity issues under other circumstances).[72]

Still, for accurately estimating potential effects, BIGS is testing further methods that have been shown to be essential for robust results. Roodman (2009) describes the process to use the GMM-Arellano-Bond on the basis of pooled Ordinary Least Squares (OLS) and the Least-Squares Dummy-Variables (LSDV) estimation that can be seen as the upper and lower bound for the estimation results of the GMM-Arellano-Bond estimator.[73] For Spain, alternatively to GMM Arellano-Bond estimator, we also use FE estimator, commonly used for panel data analysis.

[70] Y. Li et al., "A Survey on Dropout Methods and Experimental Verification in Recommendation," *IEEE Transactions on Knowledge and Data Engineering*, 2023.
[71] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021
[72] Roodman, D. (2009). 'How to do Xtabond2: An Introduction to Difference and System GMM in Stata.' *The Stata Journal*. Available at: https://doi.org/10.1177/1536867X0900900106.
[73] Ibid., p.103.

Stata allows to create files that makes the reproduction of results easy. These files (Do-Files and Log-Files) are created in the process and serve as proof of the technical work under the Behaviour Profiler & Socioeconomic Analyser.

The results from the Socioeconomic Analyser for Germany are being sent to UCSC in a static format, i.e. UCSC receives the coefficients as well as the corresponding population numbers in an Excel file so that UCSC can proceed to integrate them into the Dynamic Flows Modeler's crime estimates.

4.4 Results and Demonstration

The presentation of the results will follow in this chapter in the order of the outline of the previous sections. With the focus on preliminary and mock results from the components, in this chapter a few details need to be given as background information. A main issue that the FERMI research and analysis has encountered is availability of appropriate/compatible data to conduct a valid econometric analysis. Moreover, FERMI would require up to a certain extent access to sensitive data and the sharing thereof by the owners of that data. That is handled with a high degree of restriction and caution by authorities, not least in order to avoid non-compliance with security provisions, and potential misuse..[74] In particular, this applies to sensitive crime data, especially as far as data on politically motivated crime is concerned, which in the vast majority of cases either happens to be classified or is not assigned to the different forms of political extremism. Accordingly, the use of preliminary and mock results was embarked on as the best available option.

Nevertheless, the deliverable presents the components and first results because the processes therein are functioning. The code – if applicable – is prepared and the process of data collection is ongoing. Moreover, alternative plans of action are considered by the partners should the data considered in the beginning not be available.

4.4.1 Results of the Behaviour Profiler

The results from the Behaviour Profiler are twofold. First, one part of Equation 1 is computed by UCSC via the Dynamic Flows Modeleras a measure of likelihood of crimes occurring due to the spread of disinformation.

Second, results from the country profiles are presented. Since the whole document is too extensive to be provided for this deliverable only exemplary results are being presented briefly in sub-section Country Profiles and in more detail in the Appendix. The outcome will be made available via a separate website provided in the FERMI platform.

4.4.1.1 Crime Estimate

For demonstration purposes, subsection Crime Estimate presents a mock flow of results, from the Dynamic Flows Modeler to the Behaviour Profiler & Socioeconomic Analyser. The initial D&FN object, provided by the end-user to the platform, is firstly analysed by the Spread Analyser (T3.2) which captures its spread on social media platforms, specifically X and Mastodon. The metrics of this spread are then passed on to the Dynamic Flows Modeler, which uses them as a proxy of the intensity of the given D&FN's circulation online.

The end-user then selects a crime type and geographical location, at the NUTS-2 level, for the Dynamic Flows Modeler to estimate future occurrences, given the D&FN claim is circulating online. The Dynamic Flows Modeler calls to the Swarm Learning framework and gathers past crime occurrences in the weeks leading up to the launch of the investigation. The produced estimate is then passed on to the Behaviour Profiler

[74] See Glöckner, P. & Stuchtey, T. (2023). 'Turning Sensitive Data into Knowledge - the Need for a Common Understanding of politically Motivated Crime in Europe.' Available at: https://fighting-fake-news.eu/articles/turning-sensitive-data-knowledge-need-common-understanding-politically-motivated-crime.

& Socioeconomic Analyser which uses it in calculating economic costs to society, from the D&FN. **Fehler! Verweisquelle konnte nicht gefunden werden.** presents mock results, serving as examples for those to be generated **by the Dynamic Flows Modeler**. These mock results indicate that for a given NUTS-2 region (F11D) assaults were predicted for the subsequent 4 weeks.

Table 3: Mock predictions for a NUTS-2 region, crime type assault

Week	NUTS-2 (F11D) Assault
Week 1 (baseline)	60
Week 2	78
Week 3	60
Week 4	54
Week 5	65

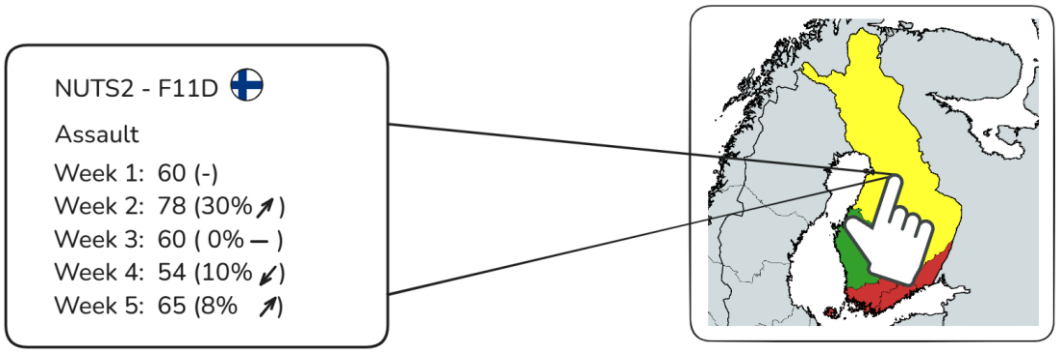

Figure 3: Example Results from Dynamic Flows Modeler

The results generated by the Behaviour Profiler are estimates of the likelihood that observed politically motivated crime is set off by disinformation spread online, taking into account other explanatory variables such as age, education and gender distribution of the population exposed to it. This result cannot be seen directly on the User Interface (UI), but it goes into the measurement of the impact which combines likelihood and severity. This, in turn, is depicted on the UI.

On top of that and as described in chapter Country Profiles, the Behavioural Profiler shows country profiles on a webpage accessible from the User Interface which comprises comparative data and information on media literacy and news consumption in various countries, as auxiliary source of information and assistance of situation assessments.

4.4.1.2 Country Profiles

The analysis conducted in D3.3 established that media literacy significantly influences societal resilience to disinformation among the FERMI target countries (Sweden, Finland, Germany, Belgium, France). With Finland and Sweden ranking highest in the European Media Literacy Index, followed by Belgium and Germany, and France ranking lowest. The data indicated a generally high awareness of disinformation, though Finnish respondents uniquely viewed disinformation as less of a national concern. Critical vulnerabilities identified include Germany's comparatively low media literacy and Belgium's fragmented media landscape and deficits in essential digital skills. The analysis supported the conclusion that media literacy can significantly help limit the spread and impact of disinformation. Nevertheless, the relationship between media

literacy and societal resilience to disinformation is multifaceted. Factors such as political climate, socioeconomic conditions, and social dynamics also substantially influence resilience. Investing in media literacy, therefore, enhances critical thinking, informed decision-making, social cohesion, and adaptability,

In this second version of the deliverable, it is apparent that trust in institutions and media plays a pivotal role in a society's resilience to disinformation. Media trust correlates strongly with disinformation resilience. Countries with higher trust in mainstream media (e.g., Finland, Sweden) are less likely to fall for false narratives, conversely, in Germany and Spain, declining trust has led to more engagement with alternative or hyper-partisan media, contributing to greater susceptibility to misinformation[75]. Countries with higher public trust and media literacy, exemplified by Finland (high trust in news, top media literacy), tend to be more resistant to disinformation, as citizens are more likely to rely on reputable sources and verify information. By contrast, where trust in government and mainstream media are low (e.g. Germany and Spain), people may be more inclined to seek news from social media or fringe outlets, making them more susceptible to conspiracy theories and propaganda.

Importantly, it has been established that these indicators are interrelated. The Media Literacy Index (which combines education, media freedom, and interpersonal trust) effectively ranks the potential resilience of societies to disinformation and fake news. Northwestern European countries with free media and educated publics top the index while also reporting higher institutional trust, indicating they are best equipped to withstand post-truth challenges. In contrast, countries scoring lower on media literacy and press freedom tend to have more polarised information environments and trust deficits, which adversaries can exploit through disinformation. Thus, bolstering media literacy and strengthening public trust in reliable institutions are critical steps in reducing susceptibility to disinformation across all five countries.

Overall, integrating these trust indicators into the media literacy analysis initiated in D3.3 and extended herein, reveals a clear pattern, where trust and media competence are strong, communities show greater collective immunity to false narratives, where trust falters, disinformation finds fertile ground. Ensuring high levels of trust in credible information sources, alongside improving media literacy education, could therefore directly enhance community resilience against the spread of disinformation.

4.4.2 Results from the Econometric Analyses of the Socioeconomic Analyser

The results of the Socioeconomic Analyser, in contrast to those of the Behavioural Profiler, do not depict the likelihood that politically motivated crimes are triggered by disinformation campaigns. Instead, they estimate the severity of the economic impact caused by a marginal increase of such crimes by one case. This represents a quantitative impact estimation, expressed in monetary terms.

This result is obtained by estimating Equation 3 for Germany and Equation 5 for Spain using the GMM Arellano-Bond estimator. The estimation yields coefficients that indicate how a one-unit increase in extremist crime affects GDP per capita, that is, economic welfare.

4.4.2.1 Results for Germany

The results for Germany are reported in Table 4 below. We estimate the impact of politically motivated crimes disaggregated by the perpetrators' motivation (left-wing, right-wing, foreign, religious) and by the nature of the crime (theft, disorderly conduct, destruction, and assault). As a result, we obtained 17 coefficients: one for each of the four crime types and four motivations, plus one for all politically motivated crimes combined. Not all coefficients are statistically significant, likely due to limited variation or low incidence of certain types of crimes in specific regions, which reduces the precision of the estimates. Thus, we use only the statistically significant coefficients.

[75] https://observatory.informationdemocracy.org/wp-content/uploads/2024/12/rapport_forum_information_democracy_2025.pdf

Type of Crime	Motivation					
				Coefficients		
All crimes	All	-0.212***				
Theft	Left		-6.670**			
Disorderly				4.011		
Destruction					0.483	
Assault						-5.498**
Theft	Right		-5.866			
Disorderly				-17.56		
Destruction					-0.407	
Assault						-3.007***
Theft	Foreign		-1.381**			
Disorderly				-1.391**		
Destruction					-1.312*	
Assault						-1.539**
Theft	Religious		17.096			
Disorderly				-30050.91		
Destruction					64.305	
Assault						-6.327*
Observations		3767	3767	3767	3767	3767
Number of Regions		401	401	401	401	401
Controls		Y	Y	Y	Y	Y
*** p<0.01, ** p<0.05, * p<0.1						

Table 4: Economic Impact of Extremist Crime in Germany

The interpretation of the coefficients is as follows: for each additional politically motivated offence, GDP per capita, measured in thousands of Euros, is reduced by the value of the respective coefficient. For example, a right-wing politically motivated assault has a coefficient of -3.007. This implies that one additional such assault per 100,000 inhabitants in a NUTS-3 region with exactly 100,000 inhabitants is estimated to reduce GDP per capita by approximately 3,000 euros over the course of a year. Similarly, the coefficient for all politically motivated crimes is -0.212. This indicates that one additional politically motivated crime per 100,000 inhabitants in such a region is estimated to reduce GDP per capita by approximately 212 Euros annually.

These coefficients are, then, sent to UCSC for the calculation of the economic cost of off-line crimes caused by online disinformation campaigns. **Table 5** reports the average cost of politically motivated crimes in each German NUTS-2 region.

Region_code	Region_name	Value
DE11	Stuttgart	-18922.1
DE12	Karlsruhe	-10003.3
DE13	Freiburg	-6810.29
DE14	Tübingen	-4109.77
DE21	Oberbayern	-19085
DE22	Niederbayern	-1823.49
DE23	Oberpfalz	-1372.94
DE24	Oberfranken	-1327.14
DE25	Mittelfranken	-3716.03
DE26	Unterfranken	-1863.04
DE27	Schwaben	-4357.47
DE30	Berlin	-27383.4
DE40	Brandenburg	-5648.4
DE50	Bremen	-1095.89
DE60	Hamburg	-4437.89
DE71	Darmstadt	-18706.1
DE72	Gießen	-1903.01
DE73	Kassel	-1922.98
DE80	Mecklenburg-Vorpommern	-2575.04
DE91	Braunschweig	-2635.03
DE92	Hannover	-5859.81
DE93	Lüneburg	-3992.56
DE94	Weser-Ems	-8779.93
DEA1	Düsseldorf	-35357.2
DEA2	Köln	-26027.6
DEA3	Münster	-10504.6
DEA4	Detmold	-5979.2
DEA5	Arnsberg	-18186.4
DEB1	Koblenz	-2940.36
DEB2	Trier	-605.225
DEB3	Rheinhessen-Pfalz	-5742.78
DEC0	Saarland	-1438.65
DEC2	Saarbrücken (approximated)	-2649.55
DED4	Chemnitz	-1961.67
DEE5	Leipzig (approximated)	-1984.12
DEF0	Schleswig-Holstein	-4086.55
DEG0	Thüringen	-9983.92
DEH0	Sachsen-Anhalt	-4021.46

Table 5: The economic cost of crime for German NUTS-2 regions

These results are summarised in an index that represent the combined outcome from severity, i.e. from the Socioeconomic Analyser, and likelihood, i.e. from the Behaviour Profiler. This index is in a speedometer on the UI visualised by a matching colour code with a range from green (low impact) to red (high impact) (see Figure 4). The impact for Germany is mid-high.

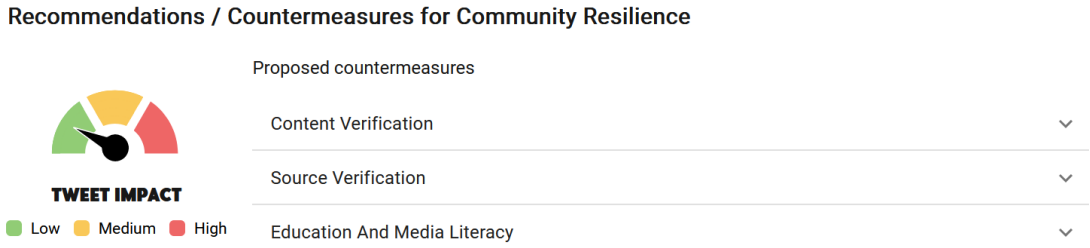

Figure 4: Visualised Output for the Behaviour Profiler & Socioeconomic Analyser

4.4.2.2 Results for Spain

The results for Spain are reported in Table 6 below. We estimated the impact of hate crimes disaggregated by the perpetrators' motivation. As a result, we obtained 12 coefficients: one for each of the 11 motivations, plus one for the total number of hate crimes. Only four coefficients are statistically significant, likely due to limited variation or low incidence of certain types of crimes in specific regions, which reduces the precision of the estimates. Therefore, in Table 6 we report only the statistically significant coefficients.

	(1)	(2)	(3)	(4)
Motivation	Coefficient			
All crimes	-0.004***			
Racism		-0.006***		
Gender discrimination			-0.052***	
Gender identity				-0.011***
Observations	408	408	408	408
Number of region_num	51	51	51	51
Controls	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Economic impact of hate crimes in Spain

These results highlight the heterogeneous economic impact of different types of hate crimes. An additional hate crime per 100,000 inhabitants is associated with a 0.4% decrease in GDP per capita, which amounts to a loss of about 120 Euros per person in a region with a GDP per capita of 30,000 euros. The effect is even more pronounced for hate crimes motivated by gender discrimination: a one-unit increase per 100,000 inhabitants leads to a 5.2% reduction in GDP per capita, corresponding to an economic loss of approximately 1,560 Euros. These findings suggest that certain forms of hate crime may impose particularly high economic costs, possibly due to their broader social repercussions and negative effects on economic participation and trust.

5 Community Resilience Management Modeler

The development of the CRMM builds on the work done in the projects ENSURESEC and IMPROVER. ENSURESEC is a project aimed to strengthen community resilience in modern e-commerce, addressing threats from malicious activities in online applications causing disruptions and fraud. By focusing on common software vulnerabilities and physical sensor interfaces, it safeguards both traditional and virtual transactions ensuring the community remains resilient in the evolving digital landscape. On the other hand, IMPROVER, enhances the resilience of European critical infrastructure by applying resilience concepts to real-life pan-European examples, including cross-border scenarios. This involves developing a methodology that integrates societal, organisational, and technological resilience concepts based on risk evaluation techniques.

The CRMM objective is to create a model that assesses the resilience level of a community in response to identified risks posed by the spread of a particular D&FN event. The component will, in short, output to the end-user (LEAs) countermeasures relevant to each initiated investigation (of disinformation-induced illegal incidents rooted in different forms of political extremism) in the FERMI platform.

The model operates by gauging the community's resilience level based on the impact of the identified risk. As the user, the LEA will receive a message indicating whether to take action, depending on the impact level of D&FN investigation. Affirmative action is recommended in the event of medium, high or extremely high impact risk; the message will include corresponding countermeasures. However, should the risk be of very low or low impact to the community, the LEA will receive no such guidance. The required "socio-economic impact assessment"[76] of the D&FN event on the community involves utilising socioeconomic and behavioural analysis, as well as crime prediction results, developed within Task 3.5 and Task 3.1.

The integration of ISO 31000 facilitates a systematic assessment of risks related to crimes associated with D&FN by receiving a calculated impact index, thus enabling effective risk treatment strategies to mitigate the impact of potential disruptions to the community through an enhanced and robust response by the LEAs. This is aligned with the principles of ISO 22300, ensuring that our development team incorporates business continuity management into decision-making processes, contributing to the overall resilience of a specific community. Equally, the adherence with the ISO 33000 underlines quality management, fostering the development of a component that not only meets regulatory requirements but also exceeds industry standards. Furthermore, ISO 27000 family for information security (see below), and specifically standard 27005, describes guidelines to help organisations establish, implement, maintain, and continually improve security management systems to ensure compliance. The component will consider the core standard contextualising within an LEA organisation as a basis for achieving and demonstrating the regulatory requirements relating to information security in the sense of actively promoting countermeasures that aid the community in recovering from the ramifications of D&FN, which particularly ensures the maintenance and continuity of security management operations.

Further specifics are provided by D4.3 and D4.4, which explain the integration of the components described in this deliverable in greater detail.

5.1 Introductory Notes on the Component

In full accordance with the GA, which stipulates that "[t]he task will start with an intense co-creation activity (that includes participatory techniques and gathering exercises, including brainstorming, focus group […])" such a focus group "performed with SSH researchers [and] practitioners"[77] was carried out in a

[76] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART B, p.5.
[77] Ibid.
The GA also alludes to the role of "citizens," which, however, seems somewhat disconnected from the objective to develop a platform for LEAs based on expert input (an ambition covering the two groups that took part in the focus group in Lisbon).

brainstorming manner in the consortium's Lisbon consortium meeting in late March 2023. The focus group addressed the rising challenges of crime related to D&FN and its dynamic. Its focus served as an initial point to understand the specific needs of each community, grasping the specificities of the proposed countermeasures and how they can be applied to the different EU communities. Another methodological focal point mentioned in the GA, namely a "DELPHI"[78] study, involves the LEAs in pointing out the countermeasures for tackling D&FN that will fill a Multi-Criteria Decision Analysis as options or alternative decisions.

The CRMM serves as valuable tool for LEAs to effectively allocate resources in employing countermeasures based on established criteria. LEAs can use the FERMI platform and, consequently, the CRMM to register investigations and to acquire adequate countermeasures to tackle disinformation events of high impact on a given community that might lead to crime. The component will support the LEA officer decision on what to do about a disinformation threat.

The CRMM is a system component belonging to the FERMI platform, responsible for the output of a set of ranked countermeasures that are in line with both the related standardisation of risk management, particularly ISO family 310xx, as well as the preferences of acting decision-makers, comprised in this case of LEAs. As explained above, it is important to note that the CRMM will only output countermeasures in the instance of a medium, high or extremely high D&FN impact, resulting in a high-risk scenario for the community; the assessment of an instance of a medium, high or extremely high impact, in turn, is informed by the Behaviour Profiler & Socioeconomic Analyser. All other instances with very low or low impact will output a script of no action to be taken by the user. The ranking of countermeasures will be produced through a value tree where on the top the user will be presented with the most recommended countermeasure and at the bottom the least recommended. The value function will take into consideration the agreed criteria and will contrast each valid option (countermeasure) with each criterion, thus making sure that the top recommended action is indeed the most agreeable. The decision-maker and end-user will be able to swiftly identify what will be the best countermeasure to tackle an instance of crime given the multitude of criteria.

The ranking will be obtained using a Multi-Criteria Decision Method (MCDM), in which the decision-makers are the various LEAs and Police Colleges (end-user partners) of FERMI from Sweden, Belgium, Germany, Finland and Spain which will be required to select a series of options or alternatives to tackle D&FN and a set of criteria, which exceeds the boundaries of the task requiring the CRMM's development, T3.3, but covers an important building block of WP4's follow-up task, T4.4 that aims at developing a "Socioeconomic D&FN Watch to enhance Community Resilience" comprising the CRMM, the Behaviour Profiler and Socioeconomic Analyser. To ensure the resulting tools are as simple and understandable as possible and to facilitate the development thereof, the partners concerned have decided to merge them right away as opposed to embarking on the rather cumbersome process of developing a CRMM that is distinct from all other components only to merge them all into the Socioeconomic D&FN Watch at a later stage. The analysis of the appropriate set of countermeasures will be done interactively by the CRMM.

Moreover, as indicated above, the accuracy of the proposed countermeasures was confirmed through the development of a DELPHI study, which aimed to produce a set of countermeasures that can be applied in each community of the associated EU partners. The method is a structured communication technique used in MCDA to gather and distil the opinions of a panel of experts or stakeholders. It is particularly valuable when there is uncertainty or lack of consensus regarding criteria weights, performance evaluations, or other aspects of the decision-making process. The criteria were proposed by BIGS and confirmed with the decision-maker through the development of a Decision Conferencing, where a smaller group of LEAs reached a consensus on which criteria to apply to our value matrix. Lastly, MCDA was be applied using a well-known method – Analytic Hierarchy Process, a decision support approach within MCDA designed to assist decision-makers in evaluating and ranking alternatives based on multiple criteria. The method structures complex problems into a hierarchy of goals, criteria, and alternatives, and uses pairwise comparisons to determine the relative importance of each element. It captures quantitative and qualitative aspects of decision problems, producing weighted scores supporting a rational and transparent decision-making process.

It is important to underline that the process of conceptualisation has been derived from a previously developed Systematic Literature Review (SLR) that followed the Kitchenham & Charters (2007) guidelines

[78] Ibid.

for literature reviews for Software Engineering.[79] The SLR aimed to find the literature terminology and consequent definition of Fake News; moreover, it aimed to identify the main concepts related to Fake News and its semantic correlations (e.g., false information, disinformation, misinformation) and further explore its relationship with the concepts of Digital Risk Management.

Secondly, the research took upon the established SLR to develop a Conceptual Modelling for mitigating high-risk Fake News. The model followed the Design Science Research methodology and consequently was demonstrated using an instance of Fake News from a public repository. Furthermore, the model was evaluated using the Bunge-Wand-Weber model to ensure its consistency and scientific value.[80] The conceptual model can be viewed in the below Figure 3 in its ArchiMate representation as it is thoroughly explained in the end of subsection 5.3.

5.2 Model Description

The proposed conceptual model is derived from a previously developed SLR where the authors identified important concepts of D&FN in the literature. In **Fehler! Verweisquelle konnte nicht gefunden werden.** below, one can see the identified concepts, their definitions according to literature and their references:

Concept	Definition
Fake News	Usually considered false or misleading information that is spread rapidly, either intentionally or unintentionally, may have drastic consequences for society, especially if widely believed; it may have the potential to polarise the different political spectrums. The SLR identified different terms present in the literature that may refer to D&FN; these include false information and misinformation. Disinformation, untrustworthy, fabricated, misleading, partisan information, digital threat, inaccurate and viral information. As laid out in D2.1 and elsewhere in this project, the disinformation notion in the sense of 1) the factual or misleading nature of the information; 2) the intention of the actors to spread such information they know to be false to obtain economic gain or deceive the public; 3) public harm caused by spreading such allegations is of interest to this research effort.[81]
Impact	These are the effects or consequences of false or misleading information on individuals or society. The impact of D&FN can vary depending on the nature of the content and the audience it reaches. The impact of D&FN can be significant and far-reaching, and it is vital to critically evaluate the information presented to prevent the spread of D&FN and its harmful effects.[82]
Context	Identifies the circumstances, events, or background information surrounding a particular story, report, and news. Understanding the context of a news item is important because it can help the reader or viewer to assess its credibility, reliability, and accuracy. Identifying the

[79] Kitchenham, B. & Charters, S. M. (2007). 'Guidelines for performing Systematic Literature Reviews in Software Engineering. Available at: https://www.researchgate.net/publication/302924724.
[80] Fettke, P. & Loos, P. (2003). 'Ontological Evaluation of Reference Models Using the Bunge-Wand-Weber Model.' *AMCIS*. Available at: https://aisel.aisnet.org/amcis2003/384.
[81] See for example (Zhang, et al. (2019)), (Habgood-Coote (2019)), (Roozenbeek und van der Linden (2019)), (Bhanu und Bhanu (2020)), (Yerlikaya und Aslan (2020)), (Ferreira, Robertson und Kirsten (2020)), (Scardigno und Mininni (2020)), (Lazar and Paun (2020)), (Alonso-García, et al. (2020)), (Flostrand, Pitt und Kietzmann (2020)), (Aswad (2020)), (Naeem, Bhatti und A. K.-H. I. & Libraries (2021)).
[82] See for example (Habgood-Coote (2019)), (Scardigno und Mininni (2020)), (Ferreira, Robertson und Kirsten (2020)), (Flostrand, Pitt und Kietzmann (2020)), (Mehta, et al. (2021)), (Petratos (2021)), (Benoliel und Becher (2022)).

	context of a particular D&FN makes it possible to understand its significance and overall implications for the community.[83]
Agent	It refers to the actor, individual, and organisations that deliberately create or spread D&FN to further their interests or agendas and possibly cause public harm too. These agents may use sophisticated techniques to make their D&FN seem more credible, such as making false allegations or creating fake images.[84] Furthermore, the *agent* is responsible for initiating the propagation of D&FN stories and is usually the main character behind the *source* (e.g., celebrities, trolls, propagandists, foreign actors, and scammers who intentionally post false content).[85]
Verifiability	The ability to check or confirm the accuracy and reliability of a piece of information. Verifiability is an essential criterion for evaluating news because it helps to determine whether the information presented is accurate, false, or misleading.
Medium	It refers to the platform, channel, or outlet that regularly produces or disseminates false or misleading information under the guise of legitimate news reporting. E.g.-Social Media Platforms, Websites, Blogs, Email Newsletter Tabloids and similar printing material, Television and News Programs, Radio Programs and Podcasts.
Event	The type of occurrence of incident related to D&FN, which can be of different types, often manipulated and misrepresented of the reality, thus creating false or misleading stories.
Source	The source of D&FN refers to the origin of the news. They originate from various sources, including individuals, organisations, and foreign governments. Again, those sources that are of interest to FERMI are the ones that intentionally spread false information to mislead the public or further their agendas.[86]
Content	Identifies the circumstances, events, or background information surrounding a particular story, report, and news. Understanding the context of a news item is important because it can help the reader or viewer to assess its credibility, reliability, and accuracy. Identifying the context of a particular D&FN makes it possible to understand its significance and overall implications for the community.[87]
Intention	It refers to the purpose or motive behind creating and disseminating false or misleading information. The intent of D&FN creators can vary and may include financial gain, the attempt to deceive the public or even cause public harm to further political propaganda, ideological or religious beliefs, and social or personal gain.[88]

[83] See for example (Naeem, Bhatti und A. K.-H. I. & Libraries (2021)), (Mehta, et al. (2021)), (Petratos (2021)), (Huber, Pospisil und Haidegger (2021)), (Marres (2018)), (Baccarella, et al. (2018)), (Habgood-Coote (2019)), (Cifuentes-Faura (2020)), (Rooke (2021)).

[84] Other strategies that fall outside the scope of the project definition are using *misleading* headlines or images or selectively presenting information.

[85] See for example (Yerlikaya und Aslan (2020)), (Ferreira, Robertson und Kirsten (2020)), (Lazar and Paun (2020)), (Petratos (2021)), (Huber, Pospisil und Haidegger (2021)), (Bastick (2021)), (Boyd (2022)).

[86] Other sources, which, however, fall outside the scope of FERMI's definition, may inadvertently share false information without realising its inaccuracies.

[87] See for example (Yerlikaya und Aslan (2020)), (Huber, Pospisil und Haidegger (2021)), (Rooke (2021)), (Macagno (2022)).

[88] See for example (Yerlikaya und Aslan (2020)), (Baccarella, et al. (2018)), (Huber, Pospisil und Haidegger (2021)), (Macagno (2022)).

Table 7: Systematic Literature Concepts

The Conceptual Model was developed using ArchiMate Full Framework in accordance to ArchiMare Specification 3.0 of The Open Group [89], where the community was modelled as an organisation or enterprise. **Fehler! Verweisquelle konnte nicht gefunden werden.** below illustrates the developed conceptual model:

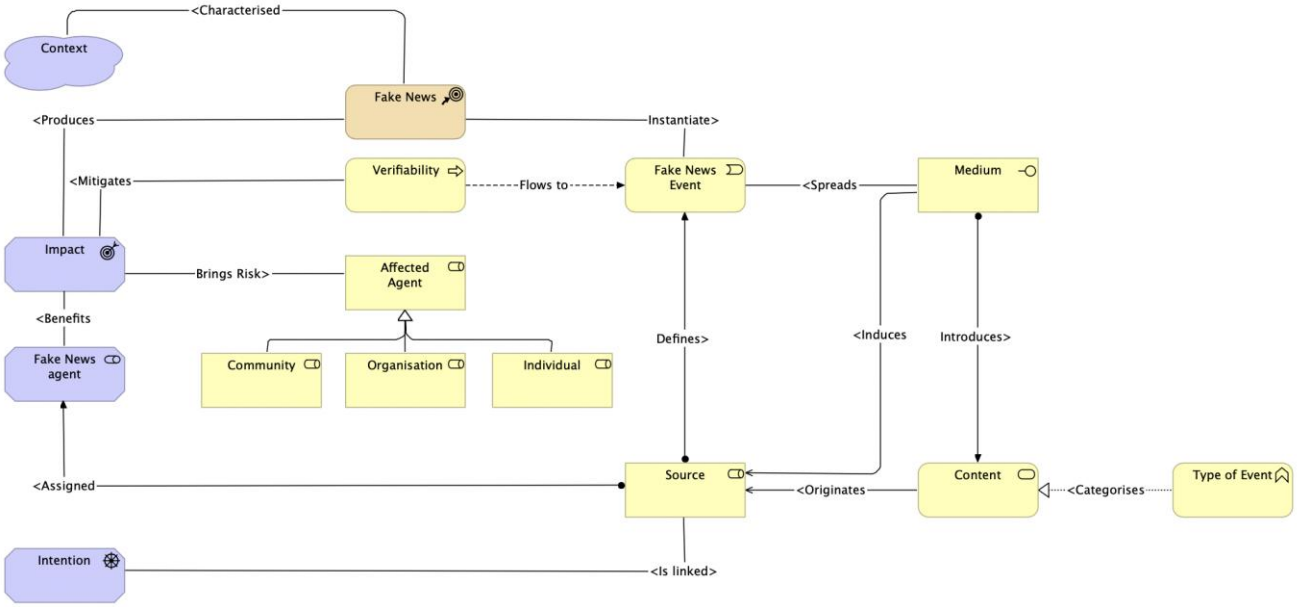

Figure 5: Conceptual Model in ArchiMate. Source: [90]

The ArchiMate modelling language was employed for constructing a model illustrating the mitigation of the impact of D&FN within a community. The conceptual model sought to depict a community as an organisation, again, in full compliance with the task description requiring such an approach,[91] making use of ArchiMate fitting for representing concepts derived from the literature, their relationships, and notations. The colour scheme patent in the model was used to differentiate between the ArchiMate layers.

Please note that given that some concepts were decomposed in order to illustrate different meanings that are presented in the scientific literature, e.g., Agent, decomposed as Fake News Agent and Affected Agent.

5.3 Technical Implementation

The technical implementation of the CRMM is informed by T3.3's further requirement that "[t]he model will be made available in a software tool, not only open source but also offered in the cloud."[92] This requires the development of a system component that is in line with the previously presented conceptual model that is adjusted to the requirements, offering scalability of performance and adherence to industry standards

[89] The Open Group, "ArchiMate® 3.1 Specification," The Open Group Standard, Document Number C197, 2019.
[90] Varela da Costa, J., Bogea Gomes, S. & Mira da Silva, M. Fake News: a conceptual model for risk management. Humanit Soc Sci Commun 11, 625 (2024). https://doi.org/10.1057/s41599-024-03096-0
[91] The GA explicitly requires that "[t]he community will be treated as an 'organization'", see 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART A, p.9.
[92] 'Grant Agreement: Project 101073980 – FERMI – HORIZON-CL3-2021-FCT-01,' *European Research Executive Agency*, 2021, PART A, p.9.

for risk management, particularly ISO family 31 XXX, which provides a framework for organisations to identify, assess, and manage risks. Its applicability can be used to manage any risk, including that of D&FN.

In compliance with the standardisation, the model identifies the risk of a particular instance that is characterised by a given context, producing a given impact that directly benefits an agent of disinformation represented in the model as Fake News Agent.

On a second layer of compliance there is the verifiability a process of verification of instances of Fake news that in essence proceed to investigate the nature of the Fake News event. The verifiability process mitigates the impact of a given event of disinformation.

It is also important to build a tool that promotes business continuity for the community complying with the standardisation 27000 family, focusing on information security management providing an adaptable framework for establishing, implementing, and continuity for organisations. To comply with this risk standardisation INOV will actively promote countermeasures that aid the community to swiftly recover from the impact of D&FN.

The CRMM is integrated into the FERMI platform through the middleware in a LINUX server solution. Initially, the Dynamic Flows Modeler interacts with INOV's component by sending a POST request to a provided endpoint containing a JSON payload. Upon receiving this input, INOV’s component processes the data using the AHP model algorithm. During this process, two separate requests are made to ChatGPT: the first to evaluate performance and the second to generate descriptive explanations for each selected countermeasure. Finally, the resulting output is sent via a POST request to a FERMI-designated endpoint.

6 Conclusions and Outlook

The present report 3.4 is an advancement and recast of deliverable D3.3 from the first period of the FERMI project, which contained the basic conceptual and methodological framework, and first mock results of the threefold component of the Behavioural Analyses and Community Resilience Facilitators package. The component provides tools that can be applied to examine the consequences of crimes occurring due to D&FN, since online activities are expected to have an impact on crime rates, more specifically, on politically motivated crime. The outputs from those sub-components and their setup as well as their integration into the FERMI platform are laid out, too. The present report presents the factual analysis conducted in the 2nd period of the project and reflects upon the results and their societal and policy consequences, and, not least, their significance for further research action.

The Behaviour Profiler updated country profiles for Spain, Germany, Belgium, Finland and Sweden. The full set of profiles will be available on the FERMI platform via a dedicated website. Media literacy is used as a core indicator of societal resilience to disinformation, with higher media-literacy levels aligning with lower susceptibility. Variation in media-consumption habits, institutional trust, and age profiles further shapes resilience to online propaganda.

In this deliverable, the country profiles have been refined with additional indicators, incorporating further data. These new insights provide a more comprehensive understanding of how trust in institutions and media literacy influence disinformation resilience. Refined media trust and literacy analysis, such as press freedom trends (RSF 2024 updates) are integrated to show changes since the previous analysis. Furthermore, impact of trust levels on susceptibility to disinformation is further elaborated, particularly with the newest OECD Drivers of Trust survey results and other relevant reports and sources.

- The findings indicate that societies with high trust and strong media literacy tend to be more resilient against disinformation-driven narratives, whereas low-trust environments present higher risks for engagement with alternative and extremist content. By integrating these societal-level risk factors, the Behaviour Profiler enhances its predictive capabilities, offering LEAs a broader framework for assessing the likelihood and impact of disinformation-induced politically motivated crimes.
- Findings on information consumption and trust: (i) It was established that the means of information consumption (TV, online news, social media, etc.) influenced media literacy and trust in news. (ii) Demographic differences were considered where data was available. (iii) Some countries showed clearer vulnerabilities due to low trust in media or high exposure to disinformation.

The Socioeconomic Analyser has established the methodology and adjusted the econometric formulas to compute a connection between D&FN induced politically motivated crime and economically measurable factors, such as GDP per capita. For this purpose, after intensive and extensive exchanges on the available criminological statistics with all FERMI LEAs, the datasets for Germany and Spain have been selected as the only ones which are fit-for-purpose to conduct the analysis. The econometric codes for both countries have been set up, considering individually the disaggregation parameters at hand, and the cost analysis has been performed considering, wherever necessary, the anonymisation/pseudonymisation of any aspects which could violate the sensitivity of those datasets. The results are elaborated on currently to achieve anonymous and applicable results that are also of practical use for the end-users of FERMI, in terms of an evidence-driven planning of resources for the sake of prevention and/or mitigation of societal impacts of disinformation on hate crime. In brief, the interpretation of the coefficients is as follows:

- In the study on Germany, for each additional politically motivated offence, GDP per capita, measured in thousands EUR, is reduced by the value of the respective coefficient. For example, a right-wing politically motivated assault has a coefficient of -3.007. This implies that one additional such assault per 100,000 inhabitants in a NUTS-3 region with exactly 100,000 inhabitants is estimated to reduce GDP per capita by approximately 3,000 EUR over the course of a year. Similarly, the coefficient for all politically motivated crimes is -0.212. This indicates that one additional politically motivated crime

per 100,000 inhabitants in such a region is estimated to reduce GDP per capita by approximately 212 EUR annually.

- The respective results for Spain highlight the heterogeneous economic impact of different types of hate crimes. An additional hate crime per 100,000 inhabitants is associated with a 0.4% decrease in GDP per capita, which amounts to a loss of about 120 euros per person in a region with a GDP per capita of 30,000 euros. The effect is even more pronounced for hate crimes motivated by gender discrimination: a one-unit increase per 100,000 inhabitants leads to a 5.2% reduction in GDP per capita, corresponding to an economic loss of approximately 1,560 euros. These findings suggest that certain forms of hate crime may impose particularly high economic costs, possibly due to their broader social repercussions and negative effects on economic participation and trust.

In both cases of Germany and Spain, the datasets have been pretty different, so the econometric models and formulas needed to be tailor-made. Nevertheless, a major insight, of direct relevance to LEAs and policymakers from planning departments form the respective ministries with the oversight on that subject matter, is the negative correlation between extremism and GDP.

As a synthesis for the overall component, not least, the Community Resilience Management Modeler has created a model that assesses the resilience level of a community in response to identified risks posed by the spread of a particular D&FN event. The component will, in short, deliver outputs to the end-user (LEAs) in the form of countermeasures relevant to each initiated investigation (of disinformation-induced illegal incidents rooted in different forms of political extremism) in the FERMI platform. The CRMM model operates by gauging the community's resilience level based on the impact of the identified risk.

In conclusion of this report, and for the sake of outlook into the future, some key takeaways from the work on this task need to be formulated. The complex issue of data availability, access, analysis, and usage is a major issue thereby. The aspect of foresight and the forward-looking security policy planning for D&FN is a complementary issue to that. Furthermore, the EU-level drafting of resilience dashboards which would need an update to include dedicated indicators on hybrid threats needs to be strengthened. Not least, the concomitant challenge of which stakeholder is mandated as an agent to decide and act, both re-actively, and, even more, pro-actively in terms of anticipatory planning on countering the negative impacts of D&FN must be addressed.

- In order to support evidence-informed security policy against the disinformation threat, be it at regional, national, or at EU-level, a new approach regarding collecting, structuring, fusing, and sharing data needs urgently to be put in place. The current state of play demonstrates unacceptable gaps, e.g. relevant data not being collected at all, and discrepancies in the way data is structured across EU Member States. Based upon the valuable insights formulated in the 2024 report[93] of the High-Level Group (HLG) on access to data for effective law enforcement to the European Commission, Member States authorities need to act upon the ways data is collected, their granularity and necessary categories of coverage, e.g. in order to consider, besides conventional crime, also emerging hybrid forms of threat, such as disinformation, which so far have not been on the radar of statistics. That said, we need to go beyond trustful inter-institutional and cross-border data-sharing and rethink the kind of relevant data that need to be collected in the future.
- Zooming out from the narrow law-enforcement perspective and considering the broader hybrid-threat challenge for security policy planning and implementation, efforts both at EU-level and in cooperation with national authorities need to strengthen the evidence basis for policies which are better fit for purpose. The EU Resilience Dashboards, introduced in 2021, assess the relative strengths and weaknesses ("capacities and vulnerabilities") for policy action of Member States along a series of composite indicators which span four dimensions: social and economic, green, digital, and geopolitical.[94] While several indicators spread along the categories of digital, geopolitical, and social & economic resilience might relate with societal and political vulnerabilities against all hybrid aspects

[93] See under https://home-affairs.ec.europa.eu/document/download/4802e306-c364-4154-835b-e986a9a49281_en?filename=Concluding%20Report%20of%20the%20HLG%20on%20access%20to%20data%20for%20effective%20law%20enforcement_en.pdf.
[94] See https://joint-research-centre.ec.europa.eu/projects-and-activities/resilience/resilience-dashboards_en.

of disinformation and foreign information manipulation and interference, there are no dedicated indicators to explicitly document the state of play with regard to its intensity and extension.

- In the same vein with the above perspective, but in a forward-looking manner, the role of foresight needs to be more strongly adopted and integrated into research activities on disinformation, and also by LEA intelligence analysts and policy planners. While evidence refers to the documentation and analysis of occurred phenomena in the past, foresight delivers insights about trends and phenomena which may happen in the future, and, in that respect, they should be treated as complementary sources for informing policy. The 2025 Joint Research Centre (JRC) report on "Emerging risks and opportunities for EU internal security stemming from new technologies" elaborates a whole array of trends and scenarios related to hybrid threats and disinformation.[95] That ought to give guidance not merely towards research avenues, but also towards LEA action.
- Last and not least, the issue of actors who need to be involved in countering the negative societal impacts, economic and other, of online disinformation and fake news, that include, of course, LEAs but also go beyond them. A valuable insight from the LEA workshop, FERMI organised in Madrid (April 2025) to demonstrate the Socioeconomic Analyser's results, was that LEAs, while central in the fight against disinformation-induced crime, do not always have the mandate to act. That is the case, e.g. under highly politicised circumstances, where public perceptions are polarised and political parties have big stakes or even instrumentalise the tense climate. Similar, when D&FN incidents have an explicit international/foreign dimension, which involve aspects of public policy besides the domestic ones. In such cases D&FN may need to be treated at a higher, political level, under inclusion of the LEAs in the exchange of evidence and the planning of the course of action.

[95] See https://publications.jrc.ec.europa.eu/repository/handle/JRC139674.

7 References

Abd El-khalek, A. M. A. 2020. "The Dynamic Relation between Population and Economic Development; a Systematic Analysis Review Considering Developing Countries' Empirical Evidence." *Journal of Economics and Finance*, 11 ed.: 12-21. https://www.academia.edu/download/62936351/C110204122120200413-95721-e1zjt.pdf.

Adjin-Tettey, Theodora Dame. 2022. "Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education." *Cogent Arts & Humanities*, 9 ed. doi:https://doi.org/10.1080/23311983.2022.2037229.

Alonso-García, S., García, G. Gómez, M. Sanz Prieto, A. J. Moreno Guerrero, and C. Rodríguez Jiménez. 2020. "The Impact of Term Fake News on the Scientific Community. Scientific Performance and Mapping in Web of Science." *Social Science*, 9 ed.: 2076-0760. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=143636290&lang=pt-pt&site=eds-live&scope=site.

Angrist, J.D., and J.-S. Pischke. 2008. "Mostly Harmless Econometrics: An Empiricist's Companion." http://diglib.globalcollege.edu.et:8080/xmlui/bitstream/handle/123456789/141/Angrist%20J.D.%2C%20Pischke%20J.-S.%20Mostly%20Harmless%20Econometrics%20%28PUP%2C%202008%29%28ISBN%200691120034X%29%28O%29%28290s%29_GL_.pdf?sequence=1&isAllowed=y.

Aswad, E.M. 2020. "In a World of 'Fake News,' What's a Social Media Platform To Do?" *Utah Law Rev*, 4 ed.: 1009–1028. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=144574921&lang=pt-pt&site=eds-live&scope=site.

Atiya, A. 2020. "Why does forecast combination work so well?" *International Journal of Forecasting.* https://doi.org/10.1016/j.ijforecast.2019.03.010.

Austin, E., B. Porismitah, and S. Doomgard. 2021. "COVID-19 disinformation and political engagement among communities of color: The role of media literacy." *Harvard Kennedy School Misinformation Review*, 1 ed. https://doi.org/10.37016/mr-2020-58.

Baccarella, C.V., T.F. Wagner, J.H. Kietzmann, and I.P. McCarthy. 2018. "Social media? It's serious! Understanding the dark side of social media." *European Management Journal*, 36 ed.: 431–438.

Bastick, Z. 2021. "Would you notice if fake news changed your behaviour? An experiment on the unconscious effects of disinformation." *Comput Human Behav*, 116 ed. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=148045922&lang=pt-pt&site=eds-live&scope=site.

Benoliel, U., and S.I. Becher. 2022. "Termination Without Explanation Contracts." *Univ Ill Law Rev*, 2022 ed.: 1059–1104. https://search.ebscohost.com/login.aspx?direct=true&db=edo&AN=157538719&lang=pt-pt&site=eds-live&scope=site.

Berthon, P., E. Treen, L. P.-N. M. I. Review, and undefined 2018. 2018. "How truthiness, fake news and post-fact endanger brands and what to do about it." *search.proquest.com.* Accessed Jan 25, 2023. https://search.proquest.com/openview/190fabaace240829d6f8ddbaabe1ae93/1?pq-origsite=gscholar&cbl=756365.

Bhanu, S., and S.M.S. Bhanu. 2020. "UbCadet: detection of compromised accounts in twitter based on user behavioural profiling." *Multimed Tools Appl*, 79 ed.: 19349–19385. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=145047986&lang=pt-pt&site=eds-live&scope=site.

BKA. n.d. *Politisch Motivierte Kriminalität (PMK) - ausländische Ideologie -*. https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKAI/PMKAI_node.html .

—. n.d. *Politisch Motivierte Kriminalität (PMK) - links -*. https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKlinks/PMKlinks_node.html.

—. n.d. *Politisch Motivierte Kriminalität (PMK) - rechts - .* https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKrechts/PMKrechts_node.html .

—. n.d. *Politisch Motivierte Kriminalität (PMK) - religiöse Ideologie -*. https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKreligioes/PMKreligioes_node.html .
—. 2024. *Politisch Motivierte Kriminalität.* https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/pmk_node.html.
BKA. 2023. "Politisch Motivierte Kriminalität im Jahr 2022." https://www.bka.de/SharedDocs/Downloads/DE/UnsereAufgaben/Deliktsbereiche/PMK/2022PMKFallzahlen.pdf?__blob=publicationFile&v=3.
BMI. 2023. "Politisch Motivierte Kriminalität im Jahr 2022." https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/nachrichten/2023/05/pmk2022-factsheets.pdf?__blob=publicationFile&v=5.
Boyd, K. 2022. "Trusting scientific experts in an online world." *Synthese*, 2021 ed.: 1–21. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=155329210&lang=pt-pt&site=eds-live&scope=site.
Bretschi, D., S. Michaelis, S. Bitschnau, and U. Kemmesis. 2023. "Kriminologische Sozialraumanalysen – Regionalräumliche Wechselwirkungen und Einflüsse sozialer Desorganisation auf politisch (rechts-)motivierte Kriminalität." In *MOTRA-Monitor 2022*, by U. Kemmesis, P. Wetzels, B. Austin, C. Büscher, A. Dessecker, S. Hutter and D. Rieger. Wiesbaden.
Bundesamt für Verfassungsschutz. n.d. *Begriff und Erscheinungsformen.* https://www.verfassungsschutz.de/DE/themen/rechtsextremismus/begriff-und-erscheinungsformen/begriff-und-erscheinungsformen_artikel.html.
Caramancion, K.M. 2021. "The Role of Subject Confidence and Historical Deception in Mis/Disinformation Vulnerability." *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2021 IEEE 12th Annual*, 0541–0546. doi:10.1109/IEMCON53756.2021.9623138.
Chiang, T.H.C., Liao, C.-S., and W.-C. Wang. 2022. "Impact of Artificial Intelligence News Source Credibility Identification System on Effectiveness of Media Literacy Education." 14 ed.: 2071-1050. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=156596850&lang=pt-pt&site=eds-live&scope=site.
Choi, Jihyang, and Jae Kook Lee. 2015. "Investigating the effects of news sharing and political interest on social media network heterogeneity." *Computers in Human Behavior*, 44 ed.: 258-266.
Cifuentes-Faura, J. 2020. "Fake news during COVID-19: how to detect them?" *Fake news durante la COVID: ¿Cómo detectarlas?*, 42 ed.: 100–103. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=145744806&lang=pt-pt&site=eds-live&scope=site.
Clark, K. A. 2005. "The Phantom Menace: Omitted Variable Bias in Econometric Research." *Conflict management and peace science*. http://www.saramitchell.org/clarke05.pdf.
Communities, Commission of the European. 2007. "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions." Brussels. https://www.cedefop.europa.eu/en/news/european-approach-media-literacy-digital-environment.
D. Mehta, D., A. Dwivedi, ,A. Patra, and M. M. Anand Kumar. 2021. "A transformer-based architecture for fake news classification." *Soc Netw Anal Min*, 11 ed.: 1–12. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=153703247&lang=pt-pt&site=eds-live&scope=site.
De Magistris, G., S. Russo, P. Roma, J.T. Starczewski, and C. Napoli. 2022. "An Explainable Fake News Detector Based on Named Entity Recognition and Stance Classification Applied to COVID-19." *Information*, 13 ed.: 2078-2489. https://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=156019008&lang=pt-pt&site=eds-live&scope=site.
de Regt, A., M. Montecchi, and S. Lord Ferguson. 2020. "A false image of health: how fake news and pseudo-facts spread in the health and beauty industry." *Journal of Product & Brand Management*, 29 ed.: 168–179.

https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=142065244&lang=pt-pt&site=eds-live&scope=site.
Eady, G., T. Pashkalis, j. Zilinsky, R. Bonneau, J. Nagler, and J.A. Tucker. 2023. “Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior.” *Nature Communications.* https://www.nature.com/articles/s41467-022-35576-9.
EC. 2023. *The impact of the Digital Services Act on digital platforms.* 3 Nov. https://digital-strategy.ec.europa.eu/en/policies/dsa-impact-platforms. .
Endtricht, Rebecca, and Janosch Kleinschnittger. 2023. "Wahrnehmung extremistischer Aktivitäten und subjektives Bedrohungserleben in Deutschland: Verbreitung und Einflussfaktoren." In *Kriminalität und Kriminologie im Zeitalter der Digitalisierung*, by L. Deyerling, A. Dreißigacker, I. Henningsmeier, M. Neumann, J. Schemmel, C. P. Schröder & L. Treskow T. Bliesener, 243-264. Mönchengladbach: Forum Verlag Godesberg GmbH.
Espaliú-Berdud, C. 2022. “Legal and criminal prosecution of disinformation in Spain in the context of the European Union.” *El Profesional de la Información*, 31 ed.: 1-14. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=157531992&lang=pt-pt&site=eds-live&scope=site.
European Parliament. 2022. *Flash Eurobarometer FL011EP: Media & News Survey 2022.* Brussels. https://data.europa.eu/data/datasets/s2832_fl011ep_eng?locale=en.
European Union. 2022. “Standard Eurobarometer STD96: Standard Eurobarometer 96 - Winter 2021-2022.” https://data.europa.eu/data/datasets/s2553_96_3_std96_eng?locale=en.
EUROSTAT. n.d. https://ec.europa.eu/eurostat/web/regions/data/database-tbd.
Farinelli, F. n.d. “Conspiracy theories and right-wing extremism – Insights and recommendations for P/CVE.” *Radicalisation Awareness Network (RAN)*, 6 ed. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwifhoaen_aDAxUi8rsIHe76CYoQFnoECBQQAQ&url=https%3A%2F%2Fhome-affairs.ec.europa.eu%2Fsystem%2Ffiles%2F2021-04%2Fran_conspiracy_theories_and_right-wing_2021_en.pdf&usg=AOvVaw1OO-H6UApi.
Ferguson, Neil, Johannes Rieckmann, and Tim H. Stuchtey. 2019. *Die Kosten des Extremismus.* 9. Vol. BIGS Standpunkt zivile Sicherheit. February. https://www.bigs-potsdam.org/app/uploads/2020/06/BIGS-Standpunkt_Nr.-9-2019_Kosten-des-Extremismus_WEB.pdf.
Ferreira, C.C., J. Robertson, and M. Kirsten. 2020. “The truth (as I see it): philosophical considerations influencing a typology of fake news.” *Journal of Product & Brand Management*, 29 ed.: 150–158. Available: https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=142065238&lang=pt-pt&site=eds-live&scope=site.
Fettke, P., and P. Loos. 2003. “Ontological Evaluation of Reference Models Using the Bunge-Wand-Weber Model.” *AMCIS.* Accessed Sep 17, 2023. https://aisel.aisnet.org/amcis2003/384.
Flostrand, A., L. Pitt, and J. Kietzmann. 2020. “Fake news and brand management: a Delphi study of impact, vulnerability and mitigation.” *Journal of Product & Brand Management*, 29 ed.: 246–254. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=142065241&lang=pt-pt&site=eds-live&scope=site.
Glöckner, P, and T. Stuchtey. 2023. “Turning Sensitive Data into Knowledge - the Need for a Common Understanding of politically Motivated Crime in Europe.” https://fighting-fake-news.eu/articles/turning-sensitive-data-knowledge-need-common-understanding-politically-motivated-crime.
Goodman, E. 2021. “Media literacy in Europe and the role of EDMO.” https://edmo.eu/wp-content/uploads/2022/02/Media-literacy-in-Europe-and-the-role-of-EDMO-Report-2021.pdf.
Grafe, Silke. 2011. "‘media literacy’ und ‘media (literacy) education’ in den USA: ein Brückenschlag über den Atlantik." *Medienpädagogik*, September 13, 20 ed.: 59-80. doi:https://doi.org/10.21240/mpaed/20/2011.09.13.X.
Gruppi, M., B.D. Horne, and S. Adalı. 2021. “NELA-GT-2020: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles.” *arXiv preprint.* https://doi.org/10.48550/arXiv.2102.04567.

—. 2022. “NELA-GT-2021: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles.” *arXiv preprint.* https://doi.org/10.48550/arXiv.2203.05659.
—. 2023. “NELA-GT-2022: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles.” *arXiv preprint.* https://doi.org/10.48550/arXiv.2203.05659.
Habgood-Coote, J. 2019. “Stop talking about fake news!” *Inquiry*, 62 ed.: 1033–1065. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=138754216&lang=pt-pt&site=eds-live&scope=site.
Harvard Kennedy School Misinformation Review. 2022. “Mapping the website and mobile app audiences of Russia’s foreign communication outlets, RT and Sputnik, across 21 countries.” 3 ed. https://doi.org/10.37016/mr-2020-110.
2023. *History of Election Fraud Claims Does Repeat Itself – and May Lead to Crimes.* https://fighting-fake-news.eu/articles/history-election-fraud-claims-does-repeat-itself-and-may-lead-crimes.
Huber, E., B. Pospisil, and W. Haidegger. 2021. “Modus Operandi in Fake News : Invited Paper.” *2021 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), Cognitive Computational Aspects of Situation Management (CogSIMA), 2021 IEEE Conference on*, 127–132. doi:10.1109/CogSIMA51574.2021.9475926.
INKAR. n.d. https://www.inkar.de/.
Islam, M.R., S. Liu, and X., Xu, G. Wang. 2020. “Deep learning for misinformation detection on online social networks: a survey and new perspectives.” *Soc Netw Anal Min*, 10 ed. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=146149934&lang=pt-pt&site=eds-live&scope=site.
Jain, D.K., A. Kumar, and A. Shrivastava. 2022. “CanarDeep: a hybrid deep neural model with mixed fusion for rumour detection in social data streams.” *Neural Comput Appl*, 34 ed.: 15129–15140. doi:10.1007/S00521-021-06743-8.
Karnowski, Veronika, Larissa Leonhard, and Anna Sophie Kümpel. 2018. "Why Users Share the News: A Theory of Reasoned Action-Based Study on the Antecedents of News-Sharing Behavior." *Communication Research Reports*, 91-100. https://anna-kuempel.de/publication/karnowski-why-2018/karnowski-et-al_2018.pdf.
Kitchenham, B., and S. M. Charters. 2007. “Guidelines for performing Systematic Literature Reviews in Software Engineering.” https://www.researchgate.net/publication/302924724.
Koehler, D. 2016. “Right-Wing Extremism and Terrorism in Europe: Current Developments and Issues for the Future.” *Prism*, 6 ed.: 84-105. https://www.jstor.org/stable/26470450.
Koohikamali, M., and A. Sidorova. 2017. “Information Re-Sharing On Social Network Sites In The Age Of Fake News.” *Inf Sci*, 20 ed.: 215–235. https://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=127807392&lang=pt-pt&site=eds-live&scope=site.
Kousika, N., Deepa. S, Deephika. C: Dhatchaiyine. B M, and J. Amrutha. 2021. “A System for Fake News Detection by using Supervised Learning Model for Social Media Contents.” *5th International Conference on Intelligent Computing and Control Systems (ICICCS), Intelligent Computing and Control Systems (ICICCS)*, 1042–1047. Accessed 2021. doi:doi: 10.1109/ICICCS51141.2021.9432096.
Lawson, M.A., and H. Kakkar. 2022. “Of pandemics, politics, and personality: The role of conscientiousness and political ideology in the sharing of fake news.” *J Exp Psychol Gen*, 151 ed.: 1154–1177. doi:10.1037/xge0001120.
Lazar, I. M., and A. C. Paun. 2020. "A predictive model for estimating citizens’ beliefs regarding the risk perception of dissemination and dispersal of fake content." *Cogn Brain Behav*, 24 ed.: 271–293. doi:doi: 10.24193/cbb.2020.24.15.
Lessinski, M. 2023. *'Bye, bye, birdie': Meeting the challenges of Disinformation. The Media Literacy Index.* Sofia: Open Society Institute.
Livingstone, S. 2004. "Media literacy and the challenge of new information and communication technologies." *Communication Review*, 1 ed.: 3-14. http://eprints.lse.ac.uk/1017.
Lynas, M. n.d. *COVID: Top 10 current conspiracy theories.* https://allianceforscience.org/blog/2020/04/covid-top-10-current-conspiracy-theories/.

Macagno, F. 2022. “Argumentation profiles and the manipulation of common ground. The arguments of populist leaders on Twitter.” *J Pragmat*, 191 ed.: 67-82. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=155851433&lang=pt-pt&site=eds-live&scope=site.
Marres, N. 2018. "Why We Can’t Have Our Facts Back." *Engag Sci Technol Soc*, 4 ed.: 423–443. doi:10.17351/ests2018.188.
Mehta, D., Dwivedi, A., A. Patra, and M. Anand Kumar. 2021. “A transformer-based architecture for fake news classification.” *Soc Netw Anal Min*, 11 ed.: 1-12. https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=153703247&lang=pt-pt&site=eds-live&scope=site.
Moore, R., and J. Hancock. 2022. “A digital media literacy intervention for older adults improves resilience to fake news.” *Scientific Reports*, 12 ed. https://www.nature.com/articles/s41598-022-08437-0.
Müller, K., and C. Schwarz. 2020. "Fanning the Flames of Hate: Social Media and Hate Crime." June 5: 1-37. https://doi.org/10.1093/jeea/jvaa045.
Murali, V. 2021. “Everything you need to know about Ensemble learning.” *Medium.*
Naeem, S., R. Bhatti, and and undefined A. K.-H. I. & Libraries. 2021. “An exploration of how fake news is taking over social media and putting public health at risk.” *Wiley Online Library*, 38 ed.: 143–149. Accessed June 2020. doi:doi: 10.1111/hir.12320.
NAMLE. 2023. *Resources-Media Literacy defined.* https://namle.net/resources/media-literacy-defined/.
Obadă, Daniel-Rares, and Dan-Christian Dabija. 2022. ""In Flow”! Why Do Users Share Fake News about Environmentally Friendly Brands on Social Media?" *Environmental Research and Public Health*, 19 ed. https://doi.org/10.3390/ijerph19084861.
OSIS. n.d. *Media Literacy Index.* Edited by Open Society Institute. Sofia. https://osis.bg/?p=4491&lang=en .
Owens, Emily, and Bocar Ba. 2021. "The Economics of Policing and Public Safety." *Journal of Economic Perspective*, 35 ed.: 3–28.
The Open Group, “ArchiMate® 3.1 Specification,” The Open Group Standard, Document Number C197, 2019.
Pedriza, Samia Benaissa. 2021. "Sources, Channels and Strategies of Disinformation in the 2020 US Election: Social Networks, Traditional Media and Political Candidates." *Journalism and Media*, 2 ed.: 605–624. https://doi.org/10.3390/journalmedia2040036.
Pennycook, Gordon, and David G. Rand. 2021. "The Psychology of Fake News." *Trends in Cognitive Sciences*, 25 ed.: 388-402. https://doi.org/10.1016/j.tics.2021.02.007.
Pérez, T., and J. Manuel. 2006. “Current trends on Media Literacy in Europe. Approaches – existing and possible – to media literacy.” https://ec.europa.eu/assets/eac/culture/library/studies/literacy-trends-report_en.pdf.
Pérez-Escoda, A., L. Pederon-Esteban, J. Rubio-Romero, and C. Jiménez-Nerros. 2021. “Fake News Reaching Young People on Social Networks: Distrust Challenging Media Literacy.” https://doi.org/10.3390/publications9020024.
Petratos, P.N. 2021. "Misinformation, disinformation, and fake news: Cyber risks to business." *Bus Horiz*, 64 ed.: 763–774. https://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S000768132100135X&lang=pt-pt&site=eds-live&scope=site.
Prskawetz, A., T. Fent, and W. Barthel. 2007. "The Relationship between Demographic Change and Economic Growth in the EU." *Vienna Institute of Demography.* https://www.oeaw.ac.at/fileadmin/subsites/Institute/VID/PDF/Publications/Forschungsberichte/FB32.pdf.
Ristea, Alina, Mohammad Al Boni, Bernd Resch, Matthew S. Gerber, and Michael Leitner. 2020. "Spatial crime distribution and prediction for sporting events using social media." *International Journal of Geographical Information Science*, 34 ed.: 1708-1739. https://doi.org/10.1080/13658816.2020.1719495.
Roodman, D. 2009. "How to do Xtabond2: An Introduction to Difference and System GMM in Stata." *The Stata Journal*, 9 ed. https://doi.org/10.1177/1536867X0900900106.

Rooke, M. 2021. “Alternative media framing of COVID-19 risks.” *Current Sociology*, 69 ed.: 584–602. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=151266773&lang=pt-pt&site=eds-live&scope=site.

Roozenbeek, J., and S. van der Linden. 2019. “The fake news game: actively inoculating against the risk of misinformation.” *J Risk Res*, 22 ed.: 570-580. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=136537868&lang=pt-pt&site=eds-live&scope=site.

Rosenbaum, P. R. 1999. “Choice as an Alternative to Control in Observational Studies.” *Statistical Science*, 14 ed.: 259-304. https://projecteuclid.org/journals/statistical-science/volume-14/issue-3/Choice-as-an-Alternative-to-Control-in-Observational-Studies/10.1214/ss/1009212410.pdf.

RSF. n.d. *Press Freedom Index.* https://rsf.org/en/index.

—. n.d. *Who are we?* https://rsf.org/en/who-are-we.

Scardigno, R., and G. Mininni. 2020. “The Rhetoric Side of Fake News: A New Weapon for Anti-Politics?” *World Futures: The Journal of General Evolution*, 76 ed.: 81-101. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=142285036&lang=pt-pt&site=eds-live&scope=site.

Schroeder, D.T., K. Pogorelov, and J. Langguth. 2019. “FACT: a Framework for Analysis and Capture of Twitter Graphs.” *019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Social Networks Analysis, Management and Security (SNAMS), 2019 Sixth International Conference on.*, 131-141. doi:10.1109/SNAMS.2019.8931870.

Silverman, Craig. 2016. *www.buzzfeednews.com.* November 16. https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook.

Terren, Ludovic, and Rosa Borge. 2021. "Echo Chambers on Social Media: A Systematic Review of the Literature." *Review of Communication Research*, 100-118.

Van Bavel, J.J., E.A. Harris, P Pärnamets, S. Rathje, K.C. Doell, and J.A. Tucker. 2022. “Political psychology in the digital (mis)information age: A model of news belief and sharing.” *Soc Issues Policy Rev*, 16 ed.: 310–311. doi:10.1111/sipr.12078.

Varela da Costa, J., Bogea Gomes, S. & Mira da Silva, M. Fake News: a conceptual model for risk management. Humanit Soc Sci Commun 11, 625 (2024). https://doi.org/10.1057/s41599-024-03096-0

Vasconcellos-Silva, P.R., and L.D. Castiel. 2020. “COVID-19, fake news, and the sleep of communicative reason producing monsters: the narrative of risks and the risks of narratives.” *Cad Saude Publica*, 36 ed. doi:10.1590/0102-311x00101920.

Viviani, M., Pasi, G. 2017. “Credibility in social media: opinions, news, and health information-a survey.” *WIREs: Data Mining & Knowledge Discovery*, 7 ed. https://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=124545337&lang=pt-pt&site=eds-live&scope=site.

W. S. Paka, W.S., R. Bansal, A. Kaushik, S. Sengupta, and T. Chakraborty. 2021. “Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection.” *Appl Soft Comput*, 107 ed. https://search.ebscohost.com/login.aspx?direct=true&db=edo&AN=150717236&lang=pt-pt&site=eds-live&scope=site.

Williams, Thomas James Vaughan, and Calli Tzani. 2022. "How does language influence the radicalisation process? A systematic review of research exploring online extremist communication and discussion." *Behavioural Sciences of Terrorism and Political Aggression.* doi:https://doi.org/10.1080/19434472.2022.2104910.

Yang, J., and Y. Tian. 2021. “Others are more vulnerable to fake news than I Am’: Third-person effect of COVID-19 fake news on social media users.” *Comput Human Behav*, 125 ed. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=152062444&lang=pt-pt&site=eds-live&scope=site.

Yerlikaya, T., and S.T. Aslan. 2020. “Social Media and Fake News in the Post-Truth Era: The Manipulation of Politics in the Election Process.” *Insight Turkey*, 22 ed.: 177–196. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=144289050&lang=pt-pt&site=eds-live&scope=site.

Zhang, C., A. Gupta, C. Kauten, A.V. Deokar, and X Qin. 2019. "Detecting fake news for reducing misinformation risks using analytics approaches." *Eur J Oper Res*, 279 ed.: 1036–1052. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=138754216&lang=pt-pt&site=eds-live&scope=site.

Appendix A Results Country Profiles

Republic of Finland: Facts[96]

- Capital: Helsinki
- Area: 338,455 sq km
- Population: 5.6 million
- Languages: Finnish, Swedish plus Sami, Karelian, Finnish Kalo

Media & Media Literacy in Finland

TV is the most popular medium in Finland however, news media are diverse. The public broadcaster Yle holds almost half of both TV and radio audiences, there are some privately owned domestic broadcasters, as well as, an increasing number of online news outlets, and several regional and local newspapers. Most media outlets are totally independent of political parties and politicians, as aforementioned, state-owned Yle is the only exception, nonetheless, politicians play no role in appointing or dismissing its journalists and reported attempts by politicians to influence media content are rare and generally not tolerated. Finland is the leading EU country in terms of press readership, and 58% of Finns say they read a press title daily [97] [98]. The state of media freedom is rated highly by international press freedom organisations[99] and more specifically Finland ranked 5th, in the 2024 Reporters Without Borders global media freedom index[100].

Main findings Finland

Finland holds a comparatively high media literacy level, and this is likely a reflection of their approach in education and their efforts in endorsing media freedom. Finland has been ranked as a world leader in education for various reasons e.g., promoting critical thinking above tests, common-sense practices and a holistic teaching environment that strives for equity over excellence[101] [102]. Finland ranks at the top of 41 European countries in their resilience to false allegations[103] [104], critical thinking and media literacy skills in Finland are initiated already at preschool across all teaching subjects.

---

[96] **Political System**: Finland is a parliamentary republic with a head of government (the prime minister) and a head of state (the president). Sauli Niinisto won the presidential election in 2012 and became the country's first conservative head of state in five decades, he won a second term in the January 2018 election. Conservative leader Petteri Orpo won a close three-way election in April 2023, defeating the centre-left incumbent Sanna Marin. He leads a coalition of the conservative NCP, nationalist Finns, the minority-language Swedish People's Party and the Christian Democrat. Finland scores consistently well on international ratings for stability, freedom, public safety and social progress. Its parliament was the first to adopt full gender equality, granting men and women the right not only to vote but also to stand for election in 1906.

[97] European Union, *Finland – EU member country profile* | European Union. Available at: https://european-union.europa.eu/principles-countries-history/country-profiles/finland_en

[98] RSF – Reporters without borders, *Finland*. Available at: https://rsf.org/en/country/finland

[99] European Union, *Finland – EU member country profile* | European Union. Available at: https://european-union.europa.eu/principles-countries-history/country-profiles/finland_en

[100] RSF - Reporters without borders, *Press freedom index - 2024*. Available at: https://rsf.org/en/index?year=2024

[101] Center for an Informed Public - University of Washington, *What we can learn from Finland*, March 01, 2023. Available at: https://www.cip.uw.edu/2023/03/01/finland-media-literacy/

[102] Mike Colagrossi, *10 reasons why Finland's education system is the best in the world*, World Economic Forum, Sept. 10, 2018. Available at: https://www.weforum.org/agenda/2018/09/10-reasons-why-finlands-education-system-is-the-best-in-the-world

[103] Marin Lessenski, *How It Started, How It is Going: Media Literacy Index 2022*, Policy Brief 57, October 2022, Open Society Institute – Sofia. Available at: https://osis.bg/?p=4243&lang=en

[104] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*.

According to the European Media Literacy Index (MLI) Report for 2023[105], Finland is placed 1st out of 41 countries in total in the respective ranking, with a score of 74 points out of 100. Finland is included in the 1st cluster of "best-performing countries", Finland is followed by Denmark, Norway, Estonia, Sweden and Ireland.

Trust in institutions and media remains comparatively high, for example, 61% of Finns trust news media, one of the highest rates in Europe, and public trust in government (47%) is above the OECD average[106]. Finland's comprehensive approach has become the go-to example for many governments and supranational bodies for addressing D&FN[107]. Evidence suggests that Finland is really invested in media literacy and the problem of fake news and disinformation, it has taken various measures and initiatives in preparing and educating citizens of all ages for the complex digital landscape of the present and the future. This suggests a strong underlying resilience. Overall, Finland's combination of top-tier education, high media trust, and press freedom makes it the most disinformation-resilient society in this cohort.

Kingdom of Sweden: Facts[108]

- Capital: Stockholm
- Area: 450,295 sq km
- Population: 10.4 million
- Languages: Swedish, plus Sami, Finnish, Meankieli, Romani, Yiddish

Media & Media Literacy in Sweden

Sweden was the world's first country to have adopted a press freedom law (1766), and it tends to hold media independence in great respect nevertheless, journalists have been targeted by threats, online hate campaigns and abusive lawsuits[109]. There are currently five media companies that own a large majority of all newspapers in the country. In the audiovisual sector the market is dominated by four television and three radio groups (including publicly owned media). Swedes have traditionally been keen newspaper readers nonetheless the press is facing pressure from digital rivals. Regarding internet-digital use, on December 2021 there were 9.6 million internet users which comprise 96% of the population[110]. TV is the most popular medium, with radio and online media following closely. Sweden ranked 3rd in the 2024 RSF Press Freedom Index but faces increasing online threats against journalists, particularly women and minority reporters. rates comparatively

Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf

[105] Ibid

[106] https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_a8004759-en/finland_596ba5da-en.html

[107] Ibid, p.3

[108] Political system – Leaders: Sweden is a constitutional monarchy and King Carl XVI Gustaf is the head of state. The monarch has primarily ceremonial and representative duties. The Head of State is the foremost representative of the Kingdom and is kept informed on issues of national importance. In accordance with representative and parliamentary democracy, the parliament enacts the laws and makes the decisions, which the government and its agencies implement. Public-private partnership is at the core of "the Swedish model", which was developed by the Social Democrats, who have governed for most of the time since the 1920s. Ulf Kristersson is the prime minster since 2022 general election and heads a coalition government between the Moderate Party, Christian Democrats and Liberals, with external support from the far-right Sweden Democrats. Relative to its population size, Sweden has taken in far more migrants and refugees than any other EU country since the beginning of the refugee crisis in 2015. As reported, this has put pressure on public services, and has prompted a rise in support for the far-right Sweden Democrats party.

[109] RSF – Reporters without borders, *Sweden*. Available at: https://rsf.org/en/country/sweden

[110] Internet World Stats – Usage and Population Statistics, *European Union Internet Users, Population and Facebook Statistics*. Available at: https://www.internetworldstats.com/stats9.htm

high for media freedom, specifically Sweden ranked fourth, after Denmark, in the 2023 RSF global media freedom index[111].

Main findings Sweden

Sweden has been promoting media literacy and online safety through formal and informal learning. The government agency "the Swedish Media Council (Statens medieråd)" is gathering, interpreting and disseminating research on children's and young people's use of media. Through this channel, information and teaching materials are produced on Media and Information Literacy (MIL) for schools and libraries. The effort is part of the EU Commission's project Safer Internet, to promote a safer and better use of the internet and mobile technologies among children and young people[112].

According to the European Media Literacy Index (MLI) Report for 2023[113], Sweden is placed 5th out of 41 countries in total in the respective ranking, with a score of 71 points out of 100. Sweden is included in the 1st cluster, "best-performing countries", where Finland is first followed by Denmark, Norway, Estonia, Sweden and Ireland[114]. Public initiatives in education and youth media literacy contribute to this high ranking. Trust in media is moderate: about 45% of Swedes trust news media, reflecting a middling level of confidence. Trust in the national government is similar (43%) and trending positively[115]. Press freedom remains very high in Sweden, and the country's legal and institutional safeguards for free expression are well-established. Thanks to these factors, Sweden shows strong resilience to disinformation. Overall, Sweden's high literacy, robust institutional trust, and media freedom create a solid foundation against disinformation, even as the government continues to address emerging challenges (e.g. hate speech and threats against journalists).

Media Literacy Analysis Belgium

Kingdom of Belgium: Facts[116]

- Capital: Brussels
- Area: 30,528 sq km
- Population: 11.5 million
- Languages: Dutch, French, German

Media & Media Literacy in Belgium

Belgium's complex institutional organization lies on both regional and linguistic grounds, with three regions, i.e. the Flemish (Flanders), Walloon (Wallonia) and Brussels (capital region) and three official languages, i.e. Dutch, French and German; the Dutch-speaking Flemish community comprises about almost 60% of the population, the French-speaking community comprises about 40% and German-speaking community just under 1% of the population (in Ostbelgien). The respective language lines also affect the educational and cultural policies, as the latter are decided at a community level. This particularity directly relates to the current division of the media markets into separate segments and inevitably increases cross-

[111] RSF - Reporters without borders, *Press freedom index - 2024*. Available at: https://rsf.org/en/country/sweden

[112] European Commission – Youth Wiki, *Sweden- Education and Training - Media literacy and safe use of new media*. Last update: 28 November 2023. Available at: https://nationalpolicies.eacea.ec.europa.eu/youthwiki/chapters/sweden/68-media-literacy-and-safe-use-of-new-media

[113] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf

[114] Ibid

[115] https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_a8004759-en/sweden_11ca1946-en.html

[116] Political System – Leaders: Belgium, located in North-western Europe, is a federal constitutional monarchy with a parliamentary system; King Philippe is the head of state, and Alexander De Croo the Prime Minister, a liberal Flemish politician who leads a seven-party coalition of liberals, socialists, greens and Christian democrats.

border influence from strong neighbouring language media (i.e. France and the Netherlands). Similarly, media literacy policies are also segmented, as they tend to address either one community or the other; each Community has its own (audiovisual) media law and a separate media regulator with sometimes varying tasks and competences[117].

The country's distinctiveness regarding its territorial and linguistic borders and their effects on the media landscape has also shaped the flow of disinformation. In particular, the issue of migration has been instrumentalized, especially by right-wing political parties; by using particular events (e.g. the terrorist attacks claimed by ISIS in 2015, the Syrian migration wave etc), the latter exacerbated the given facts in their media campaigns to reinforce their xenophobic narratives. Similarly, political parties instrumentalized the given prejudice between Dutch-speaking and French-speaking communities (also triggered by the rise of independence movements in Flanders), by spreading disinformation messages aggravating identity divisions and overall tensions. Furthermore, the COVID-19 pandemic revealed a growing presence of online communities opposing mandatory vaccination, whose online disinformation campaigns (mainly via social media platforms) were followed by violent demonstrations in Brussels against the health pass and restrictions. These communities' narratives also included messages relevant to technology scepticism, e.g. the allegedly damaging effects of 5G on health[118].

Main findings Belgium

According to the European Media Literacy Index (MLI) Report for 2023, Belgium is placed 10th out of 41 countries in total in the respective ranking, with a score of 61 points out of 100. Belgium is included in the 2nd cluster consisting in the "well-performing" countries. Comparing the Index 2023 results and the Index 2022 results, Belgium demonstrated one of the biggest improvements as it managed to move 3 positions up the ranking; it should be noted though that this was not due to the improvement of the country's actual score, as that remained the same, but rather to deterioration of the other countries' respective performances[119].

According to the 2023 Media Pluralism Monitor (MPM)[120], the country's overall scores are considered to be positive, as the efforts in the field of media literacy are continuous. However, the main two language communities demonstrate a notable difference; while the Flemish Community has had media literacy included in formal and non-formal educational settings for some time, this process is still under development in the French Community. Moreover, concerns arise regarding the market plurality, also related to the different language communities, due to the high concentration level of the media actors and the increasing market consolidation. In this regard, available information on the market share and/or ownership of digital native news media is particularly limited, as the latter are not (yet) obliged to follow the same media transparency rules imposed by their community regulators. The limited scope of action for each legislator – based on the country's institutional structure - leads to the lack of cross-media ownership restrictions, as well as lack of thresholds in media legislation.

The COVID-19 pandemic had its inevitable impact on the Belgian media landscape as well, where freelance journalists and smaller media outlets experienced the greatest losses, despite the State and Community government additional support measures. Also related to the role of the media throughout the pandemic, the report highlights an ongoing "worrisome general atmosphere of distrust and even hostility towards country's journalists in general, regardless of ideology or affiliation"[121], while particular attention is drawn to cases of online aggression against female journalists and "more notably those of ethnic background

[117] EUMEPLAT project (GA 101004488), *"D1.2: Patterns in media consumption: regional models"*, October 2021, https://www.eumeplat.eu/wp-content/uploads/2021/12/D1.2_Patterns-in-media-consumption_regional-models.pdf
[118] EU DisinfoLab / EDMO BELUX, Disinformation landscape in Belgium, May 2023, https://edmo.eu/wp-content/uploads/2023/05/20230509_BE_DisinfoFS.pdf
[119] The Media Literacy Index 2023 – Measuring Vulnerability of Societies to Disinformation, June 2023, https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf
[120] Centre for Media Pluralism and Media Freedom, *Monitoring Media Pluralism in the Digital Era: Belgium Country Report*, June 2023, Available at:https://cadmus.eui.eu/bitstream/handle/1814/75715/Belgium_results_mpm_2023_cmpf.pdf?sequence=1&isAllowed=y
[121] Ibid, p.7.

of a minority group"[122]. It further clarifies that this trend has not yet reached levels of violence, yet it points out the necessity of preventing further escalation.

The fragmented media landscape similarly affects social inclusiveness, as the efforts to guarantee inclusion of minorities and marginalised groups has proven particularly challenging. In this regard, the existing legal framework against hate speech appears to be problematic in terms of enforcement, as additional amendments to the Constitution are deemed necessary. Currently UNIA[123], the country's competent institute for monitoring and acting against discrimination, has been particularly active the past few years in prosecuting several hate speech cases in the online environment. In terms of gender equality, it should be noted that females appear to be underrepresented in management, board or CEO positions in the media industry, and similarly underrepresented in the news media, both as 'news subjects' and as 'reporters/presenters'.

Political independence of the available traditional media seems to be at a high level. Although the legal safeguards on political control and influence only apply to the broadcasting sector (radio and television), in contrast to newspapers and media distribution, the different media practitioners seem to adhere to the informal mechanisms in place maintaining their political independence, including the general constitutional protection of freedom of expression and press freedoms, and the effective self-regulatory codes of ethics for journalists. This, however, does not seem to be the case for the digital native media; considering the lack of transparency regarding ownership and/or control, affiliations to political groups cannot be confirmed, nor denied.

According to the RSF's World Press Freedom 2024 Index, Belgium is placed 16th, a significant improvement from 31st in 2023. This marks an increase in Belgium's overall press freedom score, reflecting improvements in legislative and security indicators. However, concerns remain regarding threats against journalists during protests, particularly targeting women and minority reporters. Belgium's media system remains highly concentrated, with a few dominant media families controlling the major print outlets. The public broadcasters (RTBF in Wallonia and VRT in Flanders) are politically influenced through board appointments, although they generally maintain editorial independence. Despite these concerns, Belgium has a strong tradition of press self-regulation, with media ethics overseen by the National Ethics Council[124].

Although Internet user skills in Belgium are very close to the EU average, there seem to be concerns over the high proportion of people lacking essential digital skills. In 2021, 46% of 16–74-year-olds in Belgium were classified as "digitally vulnerable", with the largest share found in Wallonia (49%), while in Flanders the share equals to 46%. Brussels residents appear to be slightly less vulnerable, while 39% are considered to be "at risk". A noteworthy remark is that the most "digitally vulnerable" people tend to be Belgians from low-income groups and those with low levels of education. This group also has the least access to so-called "essential" digital services, such as banking, health portals, government sites, or e-commerce. Furthermore, although younger people are considered to be "digitally vulnerable", they are also dealing with issues related to digital access and skills; out of the 90% of Belgians who own a phone, one in five internet users only owns a smartphone to go online.[125]

[122] Ibid, p.9.
[123] UNIA, For equality, against discrimination. Available at: https://www.unia.be/en
[124] Reporters Without Borders, *Word Press Freedom 2024 Index: Belgium Country Profile*, https://rsf.org/en/country/belgium
[125] The Brussels Times, Digital divide: Half of Belgians lack essential online skills, 2/9/2022, https://www.brusselstimes.com/282545/half-of-belgians-lack-essential-digital-skills

Federal Republic of Germany: Facts[126]

- Capital: Berlin
- Area: 357,022 sq km
- Population: 83.6 million
- Language: German

Media & Media Literacy in Germany

Media regulation is under the states, while each of the latter has its own state media law and press law. Among the few exceptions of federal regulations are the State Media Treaty (where the federal states concluded on a state treaty) and the Network Enforcement Act (federal law); these two examples display a trend in the current regulatory frameworks to increasingly take digital platforms into account. The German federal government faced challenges in the area of media and digital policy, due to disagreements within the coalition, also related to the European Commission's draft AI Regulation and the discussion on chat control. Furthermore, the Russian war of aggression in Ukraine, its social consequences and the energy crisis also affected the press sector, particularly due to the increased costs of paper and energy[127].

Germany's competitive television market is the largest in Europe (more than 38 million TV households), while the country is home to some of the world's largest media clusters. The numerous regional and national public broadcasters - organised in line with the federal political structure - compete for audiences with powerful commercial operators, while each of Germany's 16 regions regulates its own private and public broadcasting. TV networks Das Erste and ZDF, and Deutschlandradio are the national public broadcasters, while each household pays a "broadcasting contribution". Deutsche Welle (DW) is the country's international broadcaster. Newspapers and the non-tabloid press are particularly popular and, despite the many national newspapers, the press is strongest at the regional and local level. The German constitution enshrines media freedom and the press is considered to be independent; the display of swastikas and statements endorsing Nazism are illegal[128].

Main findings Germany

According to the European Media Literacy Index (MLI) Report for 2023, Germany is placed 11th out of 41 countries in total in the respective ranking, with a score of 61 points out of 100. Germany is included in the 2nd cluster consisting in the "well-performing" countries. Comparing the Index 2023 results and the Index 2022 results, Germany experienced a drop by 2 positions in the ranking (9th in 2022), accompanied by the loss of one point in the respective score (62/100 in 2022)[129].

[126] Political System – Leaders: Germany, Europe's largest economy and the most populous country in the EU, is located in the western region of central Europe and is a federal, parliamentary, representative democratic republic. The president Frank-Walter Steinmeier is the head of state, and the chancellor Olaf Scholz is the head of government since December 2021; the latter formed a coalition with the Greens and business-friendly Free Democrats, becoming the first Social Democrat chancellor since 2005, while taking over from the Christian Democrat Angela Merkel, Germany's first female chancellor, who governed for 16 years (in coalition with either the Free Democrats or the Social Democrats). The main characteristic of Germany's federal state is that responsibilities for certain subject areas are divided between the federal government and the individual states.

[127] Centre for Media Pluralism and Media Freedom, Monitoring Media Pluralism in the Digital Era: Germany Country Report, June 2023, https://cadmus.eui.eu/bitstream/handle/1814/75723/Germany_results_mpm_2023_cmpf.pdf?sequence=1&isAllowed=y

[128] BBC, Germany Media Guide, updated 28/08/2023, https://www.bbc.com/news/world-europe-17301193

[129] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf

According to the 2023 Media Pluralism Monitor (MPM)[130] the country's overall scores are considered to be satisfactory, as media diversity benefits from strong public broadcasting. However, major digital communication platforms are seen as a threat to media diversity, due to their large share of advertising budgets and their increasing role as gatekeepers for news and channels for disinformation. German lawmakers began to introduce legal regulations to counter these risks, prescribing transparency and non-discrimination for media intermediaries (e.g., social networks, search engines) and extending journalistic due diligence to all distribution channels of information services (Instagram, YouTube, etc.). However, some of these risks are still considered to be high in the digital environment, particularly in the areas of Fundamental Protection, Political Independence and Social Inclusiveness; although recent media law legislation has been focused on the digital sphere, concerns have risen that some critical issues have not been addressed, such as equal opportunities in online election campaigns and insufficient media concentration law.

Regarding the protection of Freedom of Expression, this is protected by law, and the German state does not apply arbitrary censorship on the Internet. In 2022, the "Act to Combat Right-Wing Extremism and Hate Crime" came into force, while the "Network Enforcement Act/NetzDG" (effective since 2018) obliges social networks to report certain content to the Federal Criminal Police Office if there are concrete indications relevant to criminal law that pose a threat to the democratic constitutional state, a violation of public order, child pornography, threats against life, sexual self-determination, physical integrity or personal freedom and are not justified. However, in 2022 the Cologne Administrative Court ruled that some provision of the NetzDG violate the country-of-origin principle of the E-Commerce Directive and that the Federal Office of Justice is not independent enough to monitor compliance with the obligations under the NetzDG as the competent authority. Likewise, in 2023 the OVG Münster also exempted Meta from the application of the new cross-appeal procedure on similar grounds.

In terms of Social Inclusiveness, public broadcasters in Germany currently have a clear mandate to promote integration and social cohesion. Although there are no fixed public broadcasting times for certain groups, plurality is guaranteed by the internal bodies of the broadcasters; in addition, although minorities not recognized by law do not have systematic access to public airtime, they can do so through other programs and collaborations. As regards to Gender Equality, in recent years more and more females have been appointed to management positions, while Deutsche Welle for example achieved a weighted female representation of 50%. A noteworthy remark in terms of inclusiveness is that, as the reform of public broadcasting relies on digital transformation and the conversion of certain programs to online offerings, concerns arise on whether this reform will be at the expense of the older generation, since the latter continues to watch mainly classic TV.

Media Literacy on the other hand appears to be an increasing concern in the country, although it is currently considered to be a "medium" risk. The state media authorities are responsible for promoting media literacy (which is financed by the broadcasting contribution paid by households), projects aimed at improving media literacy are being funded, while combating hate speech and disinformation are highly prioritized. However, the ground for current critique is that media competence has traditionally been assigned to the individual state media authorities as a "state matter"; nationwide issues, such as disinformation, are therefore addressed in a federal "patchwork quilt" with varying degrees of intensity. Furthermore, promoting media competence in schools can also be assigned to the area of voluntary curricula; federal differences can be found here as well, as this remains a critical issue throughout Germany. Regarding the digital environment in particular, the two digital laws, the Network Enforcement Act and the Interstate Media Treaty, address the main issues of disinformation and hate speech (including deletion obligations, journalistic due diligence obligations, and transparency and non-discrimination), yet the extent to which the laws can provide a remedy has not yet been proven.

Germany ranks 10th in the 2024 RSF Press Freedom Index, an improvement from 21st in 2023. This notable rise reflects better security conditions for journalists compared to previous years, although press freedom concerns persist, particularly regarding surveillance laws and intelligence agency powers. The legal framework for press independence is strong, but Germany continues to face SLAPP lawsuits (Strategic

[130] Centre for Media Pluralism and Media Freedom, *Monitoring Media Pluralism in the Digital Era: Germany Country Report*, June 2023, https://cadmus.eui.eu/bitstream/handle/1814/75723/Germany_results_mpm_2023_cmpf.pdf?sequence=1&isAllowed=y

Lawsuits Against Public Participation), used by powerful entities to silence investigative journalists. The economic environment for German media remains challenging, with declining print advertising revenues and increasing reliance on digital subscriptions. While public broadcasters (ARD, ZDF, Deutschlandfunk) remain highly trusted, commercial media face financial struggles[131].

Moreover, although the level of basic digital skills and basic digital content creation skills is lower than the EU average, the proportion of specialists in information and communications technology (ICT) exceeds the EU average; 49% of individuals have at least basic digital skills (EU 54%) and 65% have at least basic digital content creation skills (EU 66%). ICT specialists make up 4.9% of the labour force (compared to the EU average of 4.5%), and 4.9% of all graduates (compared to the EU average of 3.5%) are ICT graduates. The proportion of female ICT specialists is equal to the EU average, at 19%. In 2020, 24% of German businesses provided specialised ICT training for their employees. Furthermore, several reforms are incorporated into the digital transition as part of Germany's recovery and resilience plan and, regarding Digital skills, the plan includes investments in four components: providing teachers with digital devices; creating a national skills platform; creating educational competence centres; and modernising educational facilities of the national military. Investments relating to the Digitalisation of enterprises and the development and integration of digital technologies are also included in several measures.[132]

Whatsapp is the leading actively used service in Germany the past few years, were almost 84% of users confirmed this statement. Although WhatsApp is mainly a messaging service, certain features indicate similarities with social media networks, as sharing and posting between users still occurs, just not necessarily on a publicly accessible website. Facebook (61%) and Instagram (57.3%) are the next most popular platforms for the Germans[133].

Furthermore, regarding the matter of influence of online actions in the following case on twitter (now known as X), what has been reported for a particular German far-right party, Alternative für Deutschland (AfD), is its digital activism using the *hashjacking strategy (*hijacking a hashtag, defined as using someone else's hashtag to promote one's own social media activity[134]). Hashtags were designed and utilised to create a "virtual community of interested listeners" when directing users to a particular topic. As reported by Eksi [135], in the "research by the Alexander von Humboldt Institute for Internet and Society showed that the Far-right AfD supporters hijacked rising hashtags in 2020, including #FlattenTheCurve or #CoronaVirusDE (Fox, 2020). Most of the right-wing politicians use both their own party hashtags as well as the hashjacking method to strategically target opponent campaigns and to effectively polarize political discourse. As a result of their digital political communication strategy, they succeed not only online but also in elections (Darius & Stephany, 2019[136])"[137]. One can find supporting evidence to assert that parties, movements and the like, can organise and mobilise people through social media channels. Nonetheless, as analysed in a similar relevant article, "disentangling the causal relation between online activities and the public sphere is notoriously difficult"[138], what could be observed however is a "possible reinforcement process" among similar online communication

[131] Reporters Without Borders, Word Press Freedom 2024 Index: Germany Country Profile, https://rsf.org/en/country/germany

[132] Digital Skills & Jobs Platform, Germany: a snapshot of digital skills, 21/6/2023, https://digital-skills-jobs.europa.eu/en/latest/briefs/germany-snapshot-digital-skills

[133] Statista, Leading active social media and messaging platforms in Germany in 2022, https://www.statista.com/statistics/867539/top-active-social-media-platforms-in-germany/

[134] Darius, Philipp & Stephany, Fabian. (2019). "Hashjacking" the Debate: Polarisation Strategies of Germany's Political Far-Right on Twitter. 10.1007/978-3-030-34971-4_21.

[135] Sena Eksi (2022), *Digital Populism: The Internet and the Rise of Right-wing Populism,* ECPS. Available at: https://www.populismstudies.org/digital-populism-the-internet-and-the-rise-of-right-wing-populism/

[136] Darius, Philipp & Stephany, Fabian. (2019). "Hashjacking" the Debate: Polarisation Strategies of Germany's Political Far-Right on Twitter. 10.1007/978-3-030-34971-4_21.

[137] Sena Eksi (2022), *Digital Populism: The Internet and the Rise of Right-wing Populism,* ECPS. Available at: https://www.populismstudies.org/digital-populism-the-internet-and-the-rise-of-right-wing-populism/

[138] Schwemmer, C. (2021). The Limited Influence of Right-Wing Movements on Social Media User Engagement. Social Media + Society, 7(3). https://doi.org/10.1177/20563051211041650, p.10.

strategies and "reactions of the audience: more radical posts lead to more user reactions and more reactions might eventually lead to more radicalized posts by the movement."

Kingdom of Spain: Facts

• **Capital:** Madrid

• **Area:** 505,944 sq km

• **Population:** 47.6 million

• **Languages:** Spanish (Castilian) (official nationwide); co-official regional languages include Catalan, Galician, Basque, and Aranese

Media & Media Literacy in Spain

Spain's media landscape is characterised by a mix of large national media groups and strong regional media reflecting the country's linguistic diversity. Spain has four major nationwide media conglomerates (Atresmedia, Mediaset España, PRISA, and Vocento) that dominate television, radio, and print, alongside the public broadcaster RTVE. At the same time, there are influential regional outlets in Catalonia, the Basque Country, Galicia, and other communities, publishing in co-official languages and catering to regional audiences. This dual structure means the Spanish public consumes news from both national and regional sources. Press freedom is safeguarded by the 1978 Constitution, and the post-dictatorship era saw a flourishing of independent media. Journalists in Spain generally operate in a free environment, but concerns linger over issues like media ownership concentration and legal constraints (e.g. Spain's controversial "Gag Law" had provisions criticised for limiting journalism)[139]. The media market's ideological diversity mirrors Spain's political spectrum: outlets range from left-leaning to conservative, and there is a notable divide between staunchly unionist media versus those sympathetic to regional independence movements. This polarisation can influence news narratives and has occasionally been exploited to spread misleading information around sensitive topics (for instance, the Catalan independence referendum spurred waves of propaganda and misinformation on both sides). On the whole, Spain's populace is highly connected digitally, and internet penetration is high, but media literacy efforts have room to grow. The government and civil society have initiated MIL programs, yet Spain's educational curriculum has only recently begun integrating digital literacy in a systematic way[140].

Main findings Spain

In the latest European Media Literacy Index, Spain is ranked 16th out of 41 countries with a score of 58/100, placing it in the upper-middle tier of "well-performing" nations[141]. Spain benefits from relatively high educational attainment and a robust media sector, but it scores lower on interpersonal trust, a factor that can affect resilience to disinformation and fake news[142] . Public trust in media is low in Spain, only about 34% of Spaniards trust news media, one of the lowest rates in the EU. Trust in the national government is similarly low at 37%. These low trust levels suggest a vulnerability, when citizens are sceptical of traditional information sources, they may be more susceptible to misinformation or partisan narratives[143]. On the other hand, Spain's digital engagement is very high, Spaniards are active internet and social media users which both enables rapid

[139] https://thediplomatinspain.com/en/2023/05/04/spain-drops-four-places-in-rsf-world-press-freedom-index-2023/
[140] https://medialandscapes.org/country/spain
[141] Marin Lessenski, Statistical processing: Petia Brainova, Dragomira Belcheva, *"Bye, bye, birdie": Meeting the Challenges of Disinformation. The Media Literacy Index 2023. Measuring Vulnerability of Societies to Disinformation*. Policy brief. Media Literacy Index 2023 Report. Open Society Institute – Sofia. Available at: https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf
[142] https://www.aalep.eu/media-literacy-across-europe
[143] https://dadun.unav.edu/server/api/core/bitstreams/fca13c71-b8b9-4353-9d12-a8e8d9f51e14/conten

information dissemination and necessitates strong media literacy[144]. Recent surveys highlight that Spaniards increasingly get news via online platforms but also view television as a key news source[145]. In terms of disinformation resilience, Spain faces challenges from political polarization. False or misleading content has been known to thrive around politically charged issues (e.g. elections, regional autonomy debates), sometimes amplified by hyper-partisan media outlets or foreign influence campaigns[146]. Nonetheless, Spain remains among the 30% of countries that RSF deems to have a "fairly good" situation for press freedom ranked 30th, a drop compared to 36th in 2023 and the country has a favourable environment for journalism overall[147].

[144] https://en.unav.edu/web/digital-news-report/entradas/-/blogs/informe-ejecutivo
[145] https://www.statista.com/topics/10814/media-usage-in-spain/#topicOverview
[146] https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2024.1363941/full
[147] Reporters Without Borders, *Word Press Freedom 2024 Index: Spain Country Profile*, https://rsf.org/en/country/spain