FFI Norwegian Defence Research Establishment

23/02382

FFI-RAPPORT

Two layers of fog

– anonymous Norwegian websites linking to Russian-affiliated domains

Eskil Grendahl Sivertsen
Håvard Lundberg [1]
Thomas Albrechtsen [2]
Aylin Dursun [1]
Sofus Hegner [2]

[1] Analyse & Tall
[2] Common Consultancy

Two layers of fog

– anonymous Norwegian websites linking to Russian-affiliated domains

Eskil Grendahl Sivertsen
Håvard Lundberg [1]
Thomas Albrechtsen [2]
Aylin Dursun [1]
Sofus Hegner [2]

Norwegian Defence Research Establishment (FFI)
[1] Analyse & Tall
[2] Common Consultancy

6 Descember 2023

Keywords
Russland
Desinformasjon
Propaganda
Påvirkningsoperasjoner

FFI report
23/02382

Project number
1692

Electronic ISBN
978-82-464-3502-2

Approvers
Stig Rune Sellevåg, *Research Manager*
Janet Martha Blatny, *Research Director*

The document is electronically approved and therefore has no handwritten signature.

Summary

The Russian disinformation and propaganda ecosystem consists of an unknown number of ‘news’ websites in multiple languages. While some are officially Russian, such as RT and Sputnik News, the majority hide their Kremlin affiliation. We call those ‘proxy sites’. Our database includes 80 official and 2606 proxy sites. In this report, we refer to them collectively as ‘Russian-affiliated domains’.

This exploratory study examines whether there are *anonymous* Norwegian websites that share links to the 2686 Russian-affiliated domains in our database. We consider websites to be anonymous if they do not provide truthful or verifiable information about the website’s owner or authors. This *may* be an *indicator* of covert influence activity targeting a Norwegian audience and thus enable us to identify Russian-affiliated proxy sites in Norwegian.

First, we have used Majestic’s Search Engine Optimization (SEO) tool to identify domains that contain hyperlinks to the Russian-affiliated domains in our database. We have then qualitatively examined 693 identified Norwegian domains in line with common digital literacy practice to see if these domains provide credible and sufficient information to their readers about the site and its authors.

We identified 37 Norwegian domains that link to Russian-affiliated domains while operating anonymously. However, the majority of them do not seem to share Russian links today, but they appear in our dataset because previous websites on the same domain have done so in the past.

Of the 37 anonymous Norwegian domains, we have highlighted four for further analysis described in this report. These are selected not primarily because they share Russian links, but because they display various inauthentic and manipulative properties, such as authors with fake names and stolen or AI-generated portraits and plagiarized content that is auto-copied and auto-translated. Two of the four domains mimic authentic news outlets and are connected to a global network of 443 similar inauthentic and anonymous websites, of which 14 are in Norwegian. We have not found indications to suggest that these are linked to Russia, nor do we imply that they are.

While evaluating *content* is a subjective exercise, *hyperlink detection* allows us to objectively map how content from Russian-affiliated domains proliferate on the web. It provides verifiable data that can be filtered for relevance but requires further qualitative analysis. This is because outgoing hyperlinks from a domain do not provide context, for example whether as to the website supports or is making fun of Russian propaganda. One obvious limitation with this method is that we are unable to reveal domains that share content from Russian-affiliated sources if they do not link to the actual source. Another limitation is that we are only investigating hyperlinks to *known* Russian-affiliated domains.

Sammendrag

Det russiske desinformasjons- og progandanettverket består av et ukjent antall «nettaviser» på flere språk. Mens noen offisielt er russiske, som RT og Sputnik News, skjuler flertallet sin tilknytning til Kreml. Vi kaller disse «proxy-nettsteder». Databasen vår inneholder 80 offisielle nettsteder og 2606 proxy-nettsteder. I denne rapporten omtaler vi dem samlet som «russisk-affilierte domener» (*Russian-affiliated domains* på engelsk).

I denne utforskende studien undersøker vi om det finnes anonyme norske nettsteder som deler lenker til de 2686 russisk-affilierte domenene i vår database. Vi vurderer at et nettsted er anonymt hvis det ikke oppgir sannferdig eller verifiserbar informasjon om nettstedets eiere eller forfattere. Dette *kan* være en indikator på skjult påvirkningsaktivitet rettet mot et norsk publikum og kan dermed gjøre det mulig å identifisere russisk-affilierte proxy-nettsteder på norsk.

Vi har brukt Majestics Search Engine Optimization-verktøy (SEO-verktøy) for å identifisere domener som inneholder lenker til de russisk-affilierte domenene i vår database. Deretter har vi kvalitativt undersøkt 693 norske domener som vi fant på denne måten. Dette har vi gjort i tråd med alminnelige råd for kildekritikk og digital kompetanse ved å se om nettstedene på domenene gir troverdig og tilstrekkelig informasjon til leserne om eierskap og om artikkelforfattere og journalister.

Vi fant 37 norske domener som lenker til russisk-affilierte domener og hvor nettstedet på domenet opptrer anonymt. Imidlertid ser flertallet av dem ikke ut til å dele russiske lenker i dag, men fanges opp i vårt datasett fordi tidligere nettsider på samme domene har gjort det tidligere.

Av de 37 anonyme, norske nettstedene har vi fremhevet fire for ytterligere analyse, og disse er beskrevet i denne rapporten. De er valgt ikke primært fordi de deler russiske lenker, men fordi de viser tegn på inautentisk innhold og virkemidler, som artikkelforfattere med falske navn og profilbilder eller automatisk kopiering og oversettelse av innhold. To av de fire nettstedene etterligner autentiske nettaviser og er tilknyttet et globalt nettverk av 443 lignende inautentiske og anonyme nettaviser, hvorav 14 er norske. Vi har ikke funnet indikasjoner som tyder på at disse er knyttet til Russland, og vi antyder heller ikke at de er det.

Mens det å vurdere innhold er en subjektiv øvelse, gir deteksjon av lenker mulighet til å objektivt kartlegge hvordan innhold fra russisk-affilierte domener sprer seg på nettet. Metoden gir verifiserbare data som kan filtreres for relevans, men som krever ytterligere kvalitativ analyse fordi utgående lenker fra et domene ikke gir kontekst, for eksempel om nettstedet støtter eller gjør narr av russisk propaganda. En åpenbar begrensning med denne metoden er at vi ikke kan avdekke domener hvis de ikke lenker til den russisk-affilierte kilden. En annen begrensning er at vi bare undersøker lenker til kjente russisk-affilierte domener.

Contents

Preface

In a democracy, anyone is free to share links to Russian-affiliated sources. For this study, we have only tried to identify Norwegian websites that do so without providing sufficient or truthful information about site ownership and/or its authors, thus hindering people from exercising sound judgment of the websites' credibility.

While this report does not attribute any anonymous Norwegian websites to Russia or Russian affiliates, it does provide new knowledge on how hyperlinks to Russian-affiliated sites are shared in Norway and beyond. It also provides new knowledge on an unexpected area. In an exploratory study such as this, we go where the data takes us. In this case we uncovered a global network of 443 inauthentic, anonymous news sites in 32 languages, simply because two of them appeared in our dataset. We do not suggest or imply that the network has any connection to Russia.

It is our hope that this report may provide a foundation for further research, offer useful insight to Open Source analysts, journalists, digital media experts and the public, and assist in the development of open, scientifically based and transparent methods to map the Information Environment for Foreign Information Manipulation and Interference (FIMI).

Kjeller, 20 November 2023

Eskil Grendahl Sivertsen
Håvard Lundberg
Thomas Albrechtsen
Aylin Dursun
Sofus Hegner

Report at a glance

This report is based on a dataset of 2686 Russian-affiliated domains. To identify anonymous Norwegian websites that link to them, we have collected, filtered, and analysed large amounts of hyperlink data that provide new insight also beyond the purpose of this study. This is presented in this report along with the findings. Below is an overview of the key findings that form the basis of this report and are presented herein.

<table>
<tr><td colspan="4"><b>2686</b>
<b>Russian-affiliated domains in our database</b></td></tr>
<tr><td>1.46 M
domains link to the Russian-affiliated domains globally</td><td>38
percent of the 1.46 M domains are in English</td><td>ENGLISH
is the most common language for domains that share links to the Russian-affiliated domains</td><td>28
Of the Russian-affiliated domains have <i>more than</i> 100 000 domains linking to them</td></tr>
<tr><td>1356
of the 1.46 M domains that link to Russian-affiliated domains are Norwegian</td><td>693
of the Norwegian domains have online websites today</td><td>37
of the Norwegian websites are anonymous</td><td>4
of the anonymous Norwegian websites are selected for deeper analysis</td></tr>
<tr><td>10
most linked-to Russian-affiliated domains by Norwegian domains are:</td><td>1. globalresearch.ca
2. tass.com
3. english.pravda.ru
4. rt.com
5. zerohedge.com</td><td>6. sott.net
7. pravda.ru
8. sputniknews.com
9. veteranstoday.com
10. unz.com</td><td>24
of the most linked-to Russian-affiliated domains in Norway are among the 30 most linked to globally</td></tr>
</table>

1 Introduction

Russian official and affiliated websites produce vast amounts of content to create and reinforce false and misleading narratives in support of Russia's aim for reflexive control over its own population as well as foreign countries and international affairs.[1] According to the Global Engagement Center (GEC) at the U. S. State Department, the Russian disinformation and propaganda ecosystem can be described as "the collection of official, proxy, and unattributed communication channels and platforms that Russia uses to create and amplify false narratives".[2]

Proxy websites can be described as news sites, blogs and online journals that appear to be independent, legitimate and credible but are directly or indirectly affiliated with the Russian state or systematically amplify Russian propaganda and disinformation (ibid.). They play a central role in seeding and spreading Russian propaganda narratives, conspiracy theories and false or misleading news and political analyses by cross publishing each other's content.

When content from Russian-affiliated websites is shared by websites in countries outside Russia, it crosses over from the Russian disinformation and propaganda ecosystem into the domestic information environment. The information environment is here defined as "an environment comprised of the information itself; the individuals, organizations and systems that receive, process and convey the information and the cognitive, virtual and physical space in which this occurs".[3] According to the Nato Strategic Communications Centre of Excellence, "the penetration of external information influence in a domestic media ecosystem is largely enabled by the interactions between foreign and domestic actors".[4] For example, in March 2022, the Norwegian independent fact checker, Faktisk.no, described how Norwegian "alternative media" websites frequently share content from Russian-affiliated sources.[5]

In a liberal democracy such as Norway, sharing Russian propaganda and disinformation is legal and protected by the right to freedom of speech. It is up to people to decide what to believe and which sources to trust. Knowing which sources to trust has become increasingly difficult in a digital world. Navigating the information environment requires digital literacy skills. While the "alternative media" websites identified by Faktisk.no all operate in the open with verifiable owners, the Internet is full of websites that may appear credible but provide no, insufficient or incorrect information about the entity behind it. The latter is the case for many Russian-affiliated news and proxy sites in multiple languages.

The purpose of this exploratory study is to gain insight into if and how content from Russia's official state media websites (such as RT, Sputnik etc.) and proxy websites is

[1] Giles, Keir et al. (2018). *Russian Reflexive Control*. Royal Military College of Canada. Source: https://publications.gc.ca/site/fra/9.881883/publication.html

[2] Global Engagement Center. (2020). *GEC Special Report: Pillars of Russia's disinformation and propaganda ecosystem. U.S. State Department.* Source: https://www.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia%E2%80%99s-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf (p. 3).

[3] MC 0628 NATO Military policy on Strategic Communications

[4] Rodríguez, Belén Carrasco. (2020). *Information Laundering in the Nordic-Baltic region. Nato Strategic Communications Center of Excellence.* Source: https://stratcomcoe.org/publications/information-laundering-in-the-nordic-baltic-region/26

[5] Flem, S. S. & Molnes, G. (2020). Slik spres russisk propaganda i norske alternative medier. Faktisk.no. Source: https://www.faktisk.no/artikler/06epg/slik-spres-russisk-propaganda-i-norske-alternative-medier

spread to audiences in Norway via Norwegian websites that operate anonymously and/or display inauthentic properties, such as using fake identities.

There may be many reasons why someone would create anonymous Norwegian websites and use them to share links to the Russian disinformation and propaganda ecosystem. There may also be many reasons why such websites may display inauthentic properties. Neither is enough to suggest a connection to Russia in and of itself. However, such anonymous websites *may* be an indication of covert influence activity targeting a Norwegian audience and thus enable us to identify Russian-affiliated proxy sites in Norwegian. The title of this report – *Two layers of fog* – refers to the phenomenon that occurs when content from one anonymous (Russian-affiliated) website is spread by another anonymous (Norwegian) website.

This study seeks to answer the following research question(s):

1. Do Norwegian domains share links to Russian official and proxy domains in the Russian disinformation and propaganda ecosystem?

2. If so, do any of them do so anonymously?

3. If so, do any of them also display inauthentic behaviour or manipulated content?

The study is exploratory in that we first apply hyperlink detection on a large scale to identify Norwegian domains that link to Russian official and proxy domains. We then apply qualitative methods to identify which of the Norwegian domains that do so anonymously.

By "anonymously" we mean that the websites either do not reveal the owner and/or authors or operate under false identity.

This distinction is crucial. Several Norwegian domains share links to both Russian state media and proxy domains. This is legal and protected by the right to freedom of speech – one of the main pillars of democracy. The aim of this study is thus not to cast a veil of suspicion over domains that do so, but to do an *objective* and *descriptive* mapping of Norwegian websites that amplify content from Russian official or proxy sources while operating with no or little transparency.

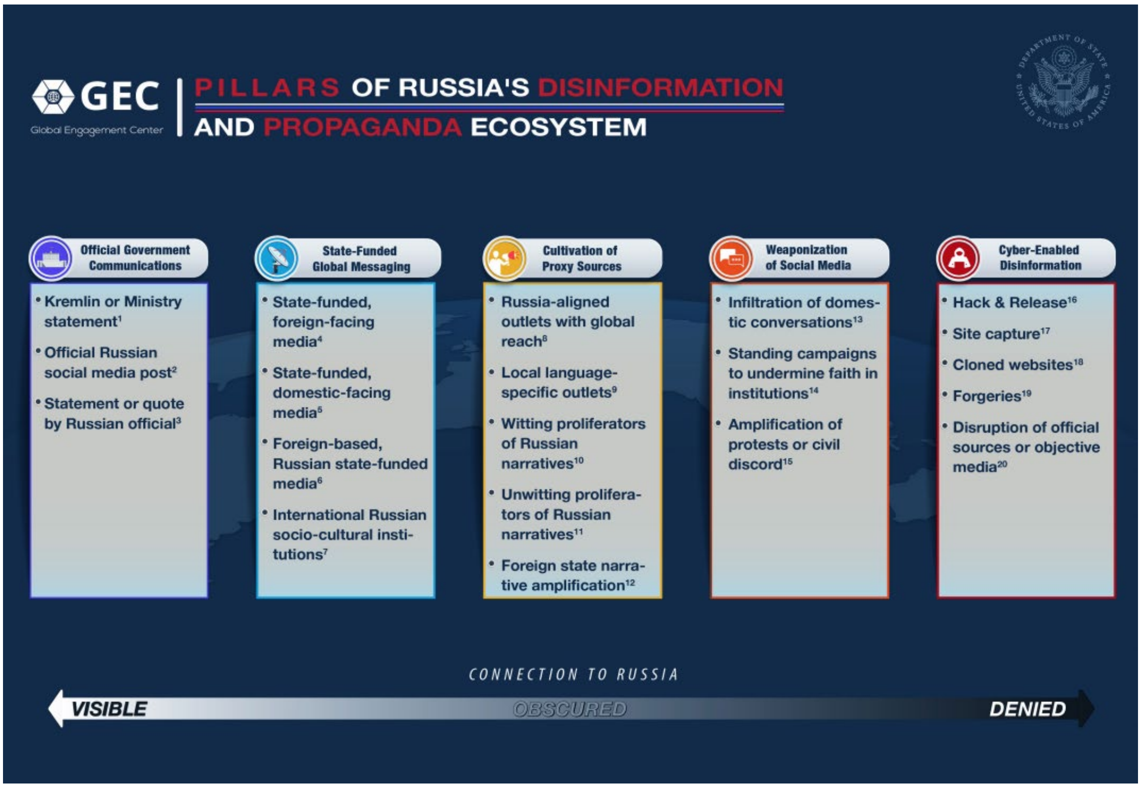
A few of the identified websites are of high interest not primarily because they have shared links to Russian-affiliated sources while operating anonymously, but because they display inauthentic properties that may shed light on various forms of manipulation techniques and other internet phenomena we believe deserve more attention. We have chosen to describe four of these websites in this report (Chapter 5 – Highlighted Norwegian domains).

2 Analytical framework

2.1 Pillars of Russia's disinformation and propaganda ecosystem

To be able to sort and separate Russian official and proxy sources in a systematic manner for data analysis, we base our approach on the logic of the U.S. State Department's Global Engagement Center's (GEC) model, *Pillars of Russia's disinformation and propaganda ecosystem.*[6] The model (Figure 2.1) seeks to provide a visual representation of the components of the ecosystem and how they work together in order to amplify and launder content across platforms.

This study includes domains in pillars 2 (i.e. Russian state media websites) and 3 (i.e. Russian affilliated proxy websites) of GEC's model. While domains in pillar 2 are officially Russian (e.g. RT, Sputnik, Tass etc.), the domains in pillar 3 appear to be independent and legitimate and hide their affiliation with Russia. Some are controlled by Russian intelligence services or other state or state-affiliated actors, and some act as systematic amplifiers of Russian propaganda.



GEC Global Engagement Center | PILLARS OF RUSSIA'S DISINFORMATION AND PROPAGANDA ECOSYSTEM

Official Government Communications
- Kremlin or Ministry statement[1]
- Official Russian social media post[2]
- Statement or quote by Russian official[3]

State-Funded Global Messaging
- State-funded, foreign-facing media[4]
- State-funded, domestic-facing media[5]
- Foreign-based, Russian state-funded media[6]
- International Russian socio-cultural institutions[7]

Cultivation of Proxy Sources
- Russia-aligned outlets with global reach[8]
- Local language-specific outlets[9]
- Witting proliferators of Russian narratives[10]
- Unwitting proliferators of Russian narratives[11]
- Foreign state narrative amplification[12]

Weaponization of Social Media
- Infiltration of domestic conversations[13]
- Standing campaigns to undermine faith in institutions[14]
- Amplification of protests or civil discord[15]

Cyber-Enabled Disinformation
- Hack & Release[16]
- Site capture[17]
- Cloned websites[18]
- Forgeries[19]
- Disruption of official sources or objective media[20]

CONNECTION TO RUSSIA

VISIBLE — OBSCURED — DENIED

Figure 2.1 Pillars of Russia's disinformation and propaganda ecosystem

[6] Global Engagement Center. (2020). *GEC Special Report: Pillars of Russia's disinformation and propaganda ecosystem. U.S. State Department.* Source: https://www.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia%E2%80%99s-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf (page 8).

2.2 Identifying Russian official and proxy domains (2nd and 3rd pillar)

We have identified 2686 domains and subdomains in pillars 2 and 3. Of these, 80 are categorized as 2nd pillar domains (Russian state media websites) and 2606 are 3rd pillar domains (Russian-affiliated proxy websites). The domains are collected from various open sources and aggregated in the table below (Table 2.1).

Table 2.1 Number of domains in our database and their source of attribution.

Source of attribution	Number of attributed domains
OpenFacto (2022)	1945
Meta (2022)	537
SMAT (2022)	54
Watts (2021), Watts (2022)	44
ODNI (2017)	38
GEC (2020), GEC (2022a)	25
ISD (2022)	14
Hooper (2017)	5
ERR (2022)	4
Lutsevych (2016)	3
EUvsDisInf (2022), EUvsDisInf (2018), EUvsDisInf (2016)	3
Marrow & Culliford (2022)	2
Vilnius Institute (2019)	2
BBC (2022)	2
Digital Forensic Research Lab (2022)	2
Agents Media (2022)	2
Van Zandt (2022a), Van Zandt (2022b)	1
Nimmo (2018)	1
Laurelle & Limonier (2021)	1
The Moscow Times (2020)	1
Debunk.org (2022)	1
CSIS & Oliker (2018)	1
Stanford Internet Observatory (2019)	1
Reuters (2017)	1

The largest batch of domains stems from OpenFacto's investigation into GRU operated domains (1945 domains), dating all the way back to the year 2000.[7] The second largest batch is from Meta's takedown of a Russian network of fake Facebook profiles and websites (537 domains) which targeted European countries in 2022.[8] [9] The batch of domains from the Global Engagement Center (GEC) contains descriptions and analyses of established and well known Russian proxy sites that play an important role in Russia's disinformation and propaganda ecosystem, including Global Research, News Front, Southfront, Geopolitica.ru, Strategic Culture Foundation, New Eastern Outlook and Katehon.[10]

3 Method

In this explorative study, we apply a two-step approach: a quantitative link-based approach to identify relevant domains and a qualitative approach to examine them. In our database we have 80 official Russian domains and 2606 Russian proxy domains that have been identified and attributed by credible third parties, including the U.S. State Department, Open Source Analysts and independent investigative journalists.[11]

First, we apply a large-scale hyperlink analysis to identify Norwegian domains that have outgoing links to the 2686 Russian-affiliated domains in our database. We then isolate domains that we define as being part of the Norwegian information environment, meaning that most of the content they provide is written in Norwegian or that they are generally directed towards a Norwegian audience (serving content through .no-domains).

The next step applies a qualitative approach to examine the websites on the identified Norwegian domains to determine whether the owner(s) and author(s) can be identified using information available on the website itself, in line with common digital literacy practices.[12] If this is not the case, we consider these domains to be anonymous.

[7] OpenFacto (2023). InfoRos' historical networks of influence. Source: https://openfacto.fr/2023/01/16/inforos-historical-networks-of-influence/

[8] Nimmo, Ben (2023). *Detailed report: Taking down coordinated inauthentic behavior from Russia and China.* Meta. Source: https://about.fb.com/wp-content/uploads/2022/11/CIB-Report_-China-Russia-Sept-2022.pdf

[9] While most, if not all, of the domains identified by Meta are inactive today, they are relevant for us in order to find out if they were linked to by Norwegian domains while they were active.

[10] Global Engagement Center. (2020). *GEC Special Report: Pillars of Russia's disinformation and propaganda ecosystem. U.S. State Department.* Source: https://www.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia%E2%80%99s-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf

[11] The complete list is not included in this report, but we make it available for other research institutions upon request. The request will be subject to approval.

[12] Medietilsynet. (2021). *Stop, think, check: How to expose fake news and misinformation*. Medietilsynet. Source: https://www.medietilsynet.no/digitale-medier/kritisk-medieforstaelse/stop-think-check-en/

Lastly, we take a closer look at the identified, anonymous Norwegian websites to see if any of them display signs of inauthentic behaviour or manipulated content, such as the use of AI generated photos, fake author profiles and AI generated or auto-translated articles. In line with the digital literacy approach, the selected websites are examined manually. Basic OSINT techniques, such as Google searches, reverse image searches and the Internet Archive Wayback Machine[13] are applied only when inauthentic content is suspected. In this report, we highlight and describe four websites because they appear inauthentic (see chapter 5 Highlighted Norwegian domains). Limitations to the methods are described in chapter 6.

3.1 Quantitative analysis - hyperlink detection

We will here describe and discuss the first, quantitative step of hyperlink analysis and the tool we have used in this study, before we describe how we applied the tool and the results it yielded.

3.1.1 About hyperlink detection and analysis

Hyperlink detection and analysis is an established practice when studying relationships between domains. Previous research has utilized this method to analyse how known disinformation sites interact with each other to spread and amplify misinformation[14], how established disinformation sites are closer connected to QAnon domains compared to established news sites[15], to identify domains linked to by known disinformation websites[16], and to analyse relationships between emerging and established media outlets in the US[17].

Compared to other methods of tracking information online, such as similar content (eg. text, images) across domains, hyperlink analysis provides some advantages.[18] Hyperlinks make the largely unstructured content of the World Wide Web more structured, as the HTML hyperlink tag (<a href=""></a>) provides a formal structure for identifying links between domains. Aggregated between domains, hyperlinks can reveal how domains are connected, implying a weaker or stronger relationship between them. As these relationships are utilized by search engines in determining search result rankings, hyperlinks have become increasingly important for search engine optimization (SEO), and in turn, in- and outgoing hyperlinks are essential when trying to increase visibility, reputation and visitor numbers online.

[13] Internet Archive Wayback Machine: https://archive.org/

[14] Sehgal, Vibhor et al. (2021). *Mutual Hyperlinking Among Misinformation Peddlers*. University of California, Berkeley. Source: https://arxiv.org/abs/2104.11694

[15] Hanley, Hans W. A. et al. (2022). *No Calm in the Storm: Investigating QAnon Website Relationships*. Stanford University. Source: https://ojs.aaai.org/index.php/ICWSM/article/view/19293/19065

[16] Waissbluth, Elliott et al. (2021). *Domain-Level Detection and Disruption of Disinformation.* University of California, Berkeley. Source: https://arxiv.org/abs/2205.03338

[17] Pak, Chankyung et al. (2020). *Intermedia Reliance and Sustainability of Emergent Media: A Large-Scale Analysis of American News Outlets' External Linking Behaviors*. International Journal of Communications. Source: https://ijoc.org/index.php/ijoc/article/view/13040

[18] Orduña-Malea, E. & Costas, R. (2021). Link-based approach to study scientifc software usage: the case of VOSviewer. Scientometrics. Source: https://link.springer.com/article/10.1007/s11192-021-04082-y

Limitations to this approach are discussed in chapter 6.

3.1.2 Utilizing the Majestic SEO tool for hyperlink analysis

To identify Norwegian domains that share links to known domains in the Russian disinformation and propaganda ecosystem, we utilize the ever-growing indexes of the Majestic Search Engine Optimization (SEO) web tool.

Majestic is one of many SEO tools available.[19] Other, similar tools include Semrush, Ahrefs, Alexa and Moz. A shared characteristic for all of them, is that they were built for SEO for website owners, and not specifically for the purpose of a study like ours, i.e. to uncover how content is interlinked online. However, because they all crawl and index content online, they provide an infrastructure for researching potential connectivity between domains. We chose Majestic because it has an impressive coverage of the web while providing an accessible and documented Application Programming Interface (API) to fetch data. Also, we have experience in applying it for this purpose through previous research.[20]

According to Majestic, it has crawled over 4.2 trillion URLs (as of June, 2023).[21] It seems to represent one of the best tools available for backlink analysis.[22] [23] [24] Recent research in the field of webometrics has also utilized Majestic.[25]

Like other SEO tools, Majestic provides an overview of backlinks to the domains they regularly crawl and index. Backlinks, also known as in-links or incoming links, is a link from some other domain to the indexed domain. Majestic utilizes a global network of web crawlers to gather and process information about webpages, including information about how domains are linked together through backlinks.

Majestic operates with what they call a 'fresh index' and a 'historic index'. The fresh index is a database containing backlink data dating approximately 3 months (90 days) back and is updated

[19] Shenoy, A., Prabhu, A. (2016). *SEO Hub: Utilities and Toolsets. In: Introducing SEO*. Apress, Berkeley, CA. Source: https://doi.org/10.1007/978-1-4842-1854-9_10

[20] Sivertsen et al. (2022). *Uønsket utenlandsk påvirkning? – kartlegging og analyse av stortingsvalget 2021.* Forsvarets forskningsinstitutt. Source: https://www.ffi.no/publikasjoner/arkiv/uonsket-utenlandsk-pavirkning-kartlegging-og-analyse-av-stortingsvalget-2021

[21] Majestic SEO tool: https://majestic.com/

[22] Jalal, S., Sutradhar, B., Sahu, K., Mukhopadhyay, P., & Biswas, S. (2015). *Search Engines and Alternative Data Sources in Webometric Research: An Exploratory Study.* DESIDOC Journal of Library & Information Technology, 35(6). Source: https://doi.org/10.14429/djlit.35.6.8883 https://publications.drdo.gov.in/ojs/index.php/djlit/article/view/8883

[23] Suad Kunosić, Denis Čeke and Enver Zerem. (2018). *Advantages and Disadvantages of the Webometrics Ranking System. In: Scientometrics Recent Advances.* IntechOpen. Source: https://www.intechopen.com/chapters/67912

[24] Varghese, M. & Lawrance, Reejo M. (2019). *Webometric Studies A Review of Literature.* In: ILIS Journal of Librarianship and Informatics Vol. 2, No. 1. pp. 91 – 101. Source: https://www.academia.edu/43766555/Webometric_Studies_A_Review_of_Literature_Reejo_M_Lawrance

[25] Dudek et al. (2021). *Co-link analysis as a monitoring tool: A webometric use case to map the web relationships of research projects.* In: Proceedings of the 18th International Conference on Scientometrics & Informetrics (2021), pp. 339-344. Source: https://arxiv.org/abs/2110.04251

daily. The historic index is a database that contains backlink data dating back to June 2006, and is updated less frequently.

3.1.3 Applying Majestic to our database of Russian-affiliated domains

We have gathered data from both the fresh and the historic index by utilizing the command *GetRefDomains* provided by Majestics API[26], using our database of Russian-affiliated domains as input. This gives us data on which, and how many, domains have linked to the Russian-affiliated domains in our database.

- The data was gathered during February 2023 (fresh in index updated 16th of February 2023, historical index updated 26th of January 2023).
- The dataset we created using Majestic's API contains approximately 1.46 million unique domains which, according to Majestic, have published hyperlinks to the known domains in the Russian disinformation and propaganda ecosystem we have in our database.
- For each of these domains, the dataset contains the aggregated number of hyperlinks from the identified domain to the domains in our database as well as information about the languages of the domains' websites.
- In total, we find that 1671 of the 2686 domains in our database have been linked to.

A domain can contain webpages with content written in different languages. Majestic provides information about the different detected languages for the webpages belonging to the domain, the confidence level of the detected languages and the proportion of content in each language. When determining the language of a singular web page, Majestic considers the content of the anchor text of the backlink and the page title of the page on which the hyperlink appears. It then aggregates the number of pages with identified languages for the whole domain, which is the variable we have used to identify Norwegian domains. The language information is based on Majestic's own internally developed algorithms, which at the time of writing was not documented online.

3.1.4 Processing and filtering data

To identify the domains relevant for this study, we applied a step-by-step filtering technique. This is described in the following paragraphs and visually illustrated in Figure 3.1 – Visual representation of the step-by-step filtering process.

Step 1
As the goal of our hyperlink analysis is to identify Norwegian websites that share links to domains in Russia's disinformation and propaganda ecosystem, we used Majestic's detected language information to filter down the returned list of **1.46 million domains**. By filtering out

[26] https://developer-support.majestic.com/api/commands/get-ref-domains.shtml

all the domains that are *not* .no domains or have content written in Norwegian (bokmål or nynorsk – the two official Norwegian versions), we end up with **9751 domains**. Sámi languages were not included as they were outside the scope of this project.

Step 2

To account for levels of uncertainty in the precision of language detection, we applied two more criteria to increase the accuracy of our findings:

1) The content of the domain (i.e. the webpages on it) must at least contain 50% Norwegian language (bokmål or nynorsk)

   and

2) Majestic's confidence level of the detected language must be 70% or higher.

This combination of filters gives us **1356 domains**.

Step 3

To be included for further analysis, we filter out all domains which have less than 3 identified outgoing links to our proxy sites. The cut-off value was set to limit the dataset to a manageable size while still including relevant websites for further research. This leaves **976 domains**.

Step 4

We then access these domains to see if they are still accessible, as Majestic's historical index can contain inactive domains. We acquire the returned HTTP response status codes for all domains (ranging from 100-599), and we filter out the domains which do not return a HTTP response (i.e. connection error). This leaves us with a shortlist of **693 domains** for further, qualitative analysis to see how many of these are anonymous and if any of them display inauthentic properties.

As the visual representation in Figure 3.1 shows, the last, qualitative analysis yields 37 domains. The analysis leading from 693 to 37 domains is described in the following section, 3.2 – Qualitative analysis.

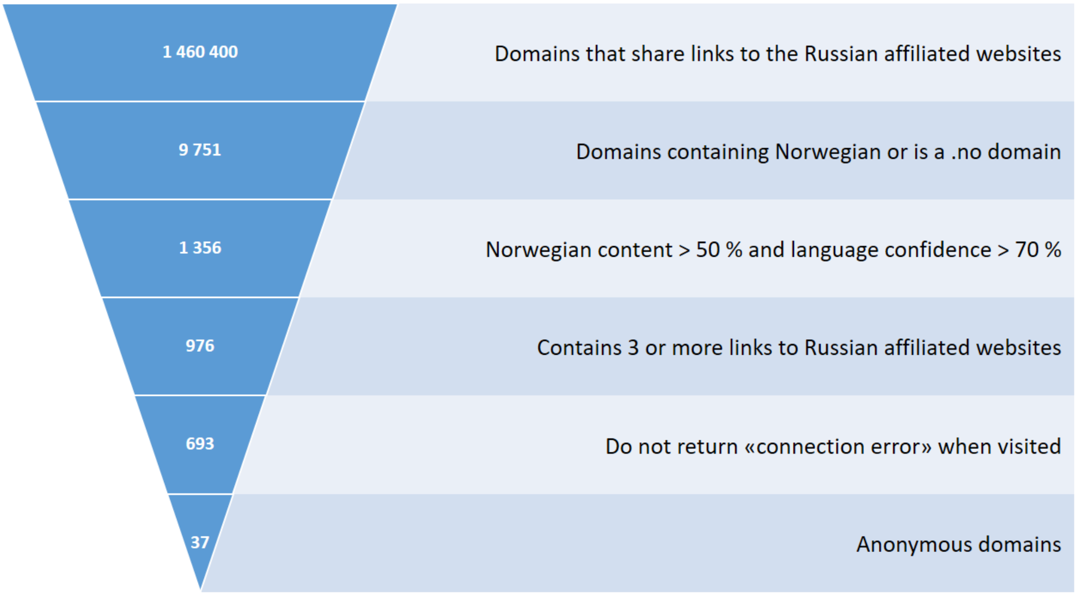


Figure 3.1 Visual representation of the step-by-step filtering process

3.2 Qualitative analysis

The next step of our analysis is to find out whether the remaining 693 domains can be attributed to verifiable owners or authors. To do so, we manually screen all 693 domains to see if they provide verifiable information about site ownership and/or its author(s).

The steps of this process are kept simple, as we approach this from a digital literacy standpoint; the information should be easily visible and accessible to regular internet users and not require Open Source Intelligence (OSINT) skills or tools.

- We search for the relevant information by browsing the websites like a regular internet user would, in line with common digital literacy practices and common sense.

- We only assess information that is currently publicly available and visible on the specific domain (i.e. the website).

- When in doubt, we perform simple Google searches to see if we can verify owner or author names or the email addresses provided by the website as contact information.

- We consider the website verified if it provides credible information about the website itself, such as names, photos and descriptions of the owner and/or authors, or information such as addresses, telephone numbers and/or VAT-number (Norwegian “organisasjonsnummer”), i.e. identity markers that you should expect to find on a legitimate website and that internet users are advised to look for when evaluating trustworthiness. Of course, this means that a potential covert actor that has made substantial efforts to come across as legitimate may be overlooked.

4 Results

To sum up, out of the 1 460 400 domains that link to Russian-affiliated domains, 9 751 contain Norwegian language or are .no domains. When removing domains that have less than 50% percent Norwegian content and a language detection confidence lower than 70%, we are left with 1 356 domains. Of these, 976 share three or more links to Russian-affiliated domains. 693 of them did not return a connection error when visited. In conclusion, we identified 693 active Norwegian domains that share three or more links to Russian-affiliated source.

Of the 693 identified, active Norwegian domains that share links to Russian-affiliated domains, 151 are discarded because the information on the website is not accessible. The website on 500 domains are categorized as “verifiable” based on information about site ownership and/or its authors found on the website. Five websites are categorized as “partly verifiable” (see definition in section 4.1.1 Partly verifiable websites). 37 websites are categorized as “anonymous”. The websites of interest to this report are those that are categorized as “anonymous” and “partly verifiable”.

4.1.1 Partly verifiable websites

We consider 5 of the 693 websites partly verifiable. Because we have evaluated the websites at face value only based on common digital literacy practices, some sites fall into a grey zone in which we cannot conclude that the website operates anonymously, nor can we conclude that it doesn’t. These sites display properties such as:

- Provides a credible and, in a Norwegian context, unique name of the owner, but with no or non-verifiable contact information.
- Does not state the owner on the domain itself and features many anonymously written articles, but has a name connected to affiliated websites and some articles under a byline with the same name, that is verifiable.
- Reveals no owner, but features what seems to be authentic, Norwegian contributors.

We have not analysed these five partly verifiable websites further. This is due to privacy considerations. As FFI is a governmental research institution, we have set the bar high to avoid risking a situation in which we could potentially cast a veil of suspicion over authentic Norwegian citizens’ legal activities online.

4.1.2 Anonymous websites

We consider 37 of the 693 websites anonymous because they provide no or insufficient information about site ownership or authors of articles. These websites only have a generic contact form or email address (in the format of “info@nameofdomain.com”) and/or provide very common or generic author names (such as "Jonas Jørgen") or authors named after celebrities (such as "Edvard Munch) with no contact details or other verifiable information.

Several of the anonymous and the partly verifiable websites seem to appear in our dataset not because they have shared links to Russian sources recently, but because an older website on the same domain has in the past (see "Limitations" in chapter 6). Therefore, we choose not to list all the anonymous websites in this report. Instead, we have selected four out of the 37 anonymous websites for further analysis in this report. They are highlighted not because they share many links to Russian-affiliated domains, but because they display inauthentic properties and manipulation techniques we believe should be made public knowledge in order to increase people's digital literacy abilities (see chapter 5 - Highlighted domains).

4.1.3 Websites with inauthentic properties

The four highlighted, anonymous websites display inauthentic behaviour and/or employ manipulation techniques. These are described in chapter 5 - Highlighted domains. Examples include auto-copying and auto-translation of content taken from authentic news sites, posing as legitimate Norwegian news sites while being part of a global, inauthentic network, and using fake identities of editorial staff and artificially generated profile pictures (so-called GAN images).

4.1.4 Other relevant findings regarding the spread of Russian propaganda

While the purpose of this study has been to identify anonymous Norwegian websites that share links to websites in the Russian disinformation and propaganda ecosystem, our analysis revealed other interesting findings related to the spread of content from Russian-affiliated sources in general. While these findings have not been subject for further analysis in this report, we describe them here for the purpose of sharing information that may add to existing knowledge of the reach and nature of the network of Russian-affiliated domains.

4.1.4.1 Top ten languages for sharing links to Russian-affiliated domains

Our research with Majestic revealed that 1.46 million domains globally contain hyperlinks to the 2686 domains in our dataset of known domains in the Russian disinformation and propaganda ecosystem. Of these 1.46 million domains, the majority are in English (38%), followed by Russian (29%). The third biggest language is Chinese (5%), followed by Spanish (4%), German (2%), French (2%), Korean (2%), Japanese (2%), Arabic (2%) and Italian (1%). (Figure 4.1). The data may be indicative of the prioritization of languages, and thus target groups, for Russian influence and propaganda.

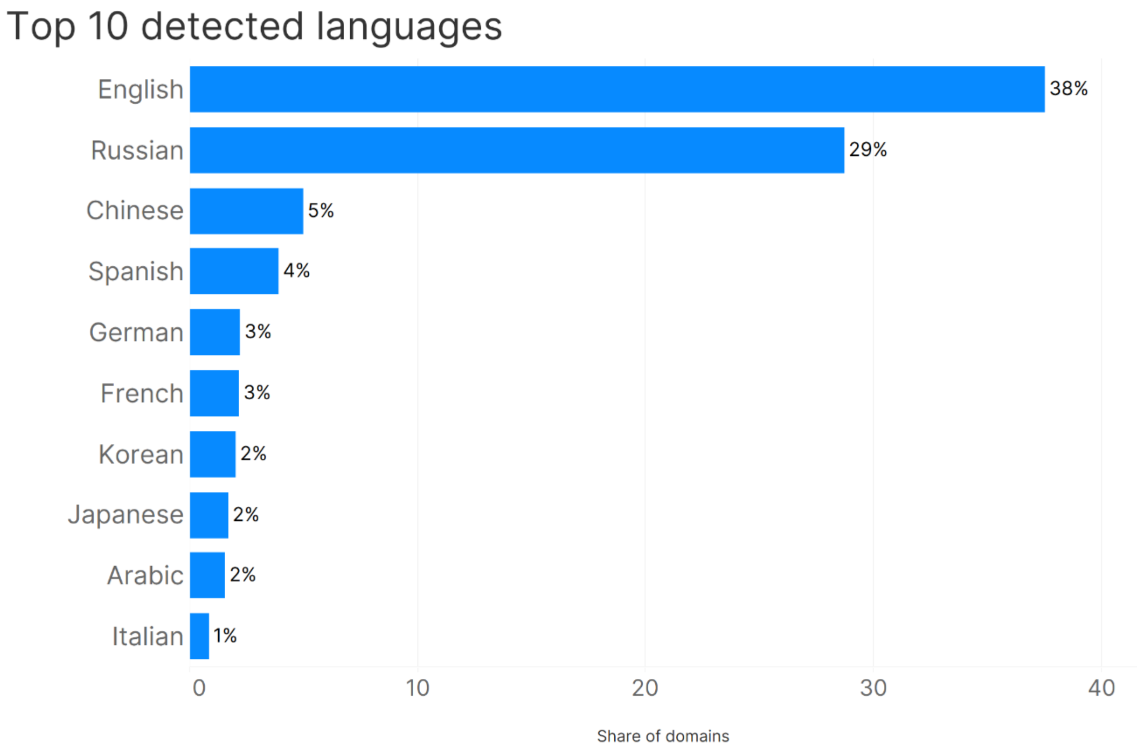


Figure 4.1 The 10 largest detected languages among the 1.46 million domains that have shared links to Russian-affiliated sources. Note that each domain has been assigned to the main language detected on the domain. Beyond the top ten languages presented in this figure, 6% account for 116 languages and 7% did not have a detected language.

4.1.4.2 Link sharing in Nordic languages

Only 0,11% of the detected 1.46 million domains are Norwegian. Swedish is the largest of the Nordic languages (0.20%), followed by Danish (0.13%) and Finnish (0.12%). At the bottom of the list we find Icelandic (0.01%) and Faeroese (0.002%). (Figure 4.2). This seems to reflect the population size of the Nordic countries.

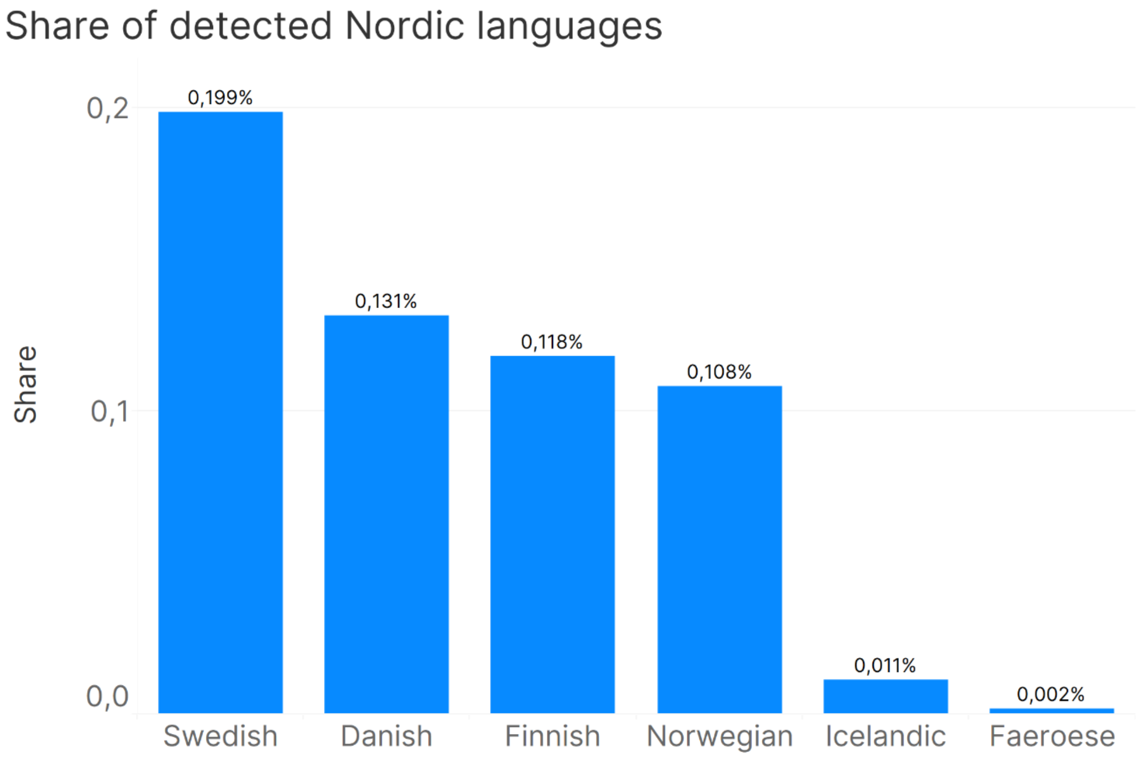


Figure 4.2 Share of Nordic languages in the dataset of 1.46 million domains that have shared links to Russian-affiliated domains. Note that each domain has been assigned to the main language detected on the domain.

4.1.4.3 The most linked to Russian-affiliated domains globally

Of the 30 most linked-to Russian-affiliated domains in our dataset, 28 have more than 100 000 back-linking domains ("referring domains"), going above Majestic's cut-off threshold. This means that there are at least 28 domains in the Russian disinformation and propaganda ecosystem that have *more than* 100 000 domains linking to them. We are unable to identify how many, as Majestic only provides exact numbers up to its 100 000 cut-off value. The domains are presented in Table 4.1.

Table 4.1 The top 30 most linked-to domains in our database of domains in the Russian disinformation and propaganda ecosystem. The top 28 have more backlinks than Majestic's maximum cut-off value of 100 000.

Domain in the Russian ecosystem	Number of domains linking to it globally
vesti.ru	More than 100 000
rbc.ru	More than 100 000
inosmi.ru	More than 100 000
zerohedge.com	More than 100 000
unz.com	More than 100 000
globalresearch.ca	More than 100 000
voltairenet.org	More than 100 000
english.pravda.ru	More than 100 000
regnum.ru	More than 100 000
veteranstoday.com	More than 100 000
sott.net	More than 100 000
rbth.com	More than 100 000
sputniknews.com	More than 100 000
1tv.ru	More than 100 000
tass.com	More than 100 000
ntv.ru	More than 100 000
rt.com	More than 100 000
tass.ru	More than 100 000
ria.ru	More than 100 000
interfax.ru	More than 100 000
gazeta.ru	More than 100 000
izvestia.ru	More than 100 000
kp.ru	More than 100 000
life.ru	More than 100 000
vedomosti.ru	More than 100 000
pravda.ru	More than 100 000
rg.ru	More than 100 000

lenta.ru	More than 100 000
tvzvezda.ru	95 954
russian.rt.com	90 374

4.1.4.4 The most linked to Russian-affiliated domains by Norwegian domains

When reviewing *all* the identified 1356 Norwegian domains (.no domains and domains in Norwegian language with high confidence level), we find that 24 out of the top 30 most linked-to Russian-affiliated domains are also among the top 30 linked-to domains overall (globally). In addition, the following 6 domains are not part of the overall top 30 domains: *from-ua.com, strategic-culture.org, russia-insider.com, southfront.org, site.ru,* and *thesaker.is*. We have not looked into why these seem to be more popular to backlink to in a Norwegian context, as that is beyond the scope of this study. Here is the list of the Russian-affiliated domains that are most linked to by Norwegian domains (Table 4.2).

*Table 4.2 The top 30 most linked-to Russian-affiliated domains from Norwegian domains. 24 of them are on the top 30 list overall. The six domains that are on the Norwegian top 30 list, but not on the global top 30 list are marked with the asterisk *.*

Domain in the Russian ecosystem	**Number of Norwegian domains linking to it**
globalresearch.ca	187
tass.com	151
ussian.pravda.ru	141
rt.com	137
zerohedge.com	136
sott.net	127
ussia.ru	113
sputniknews.com	103
veteranstoday.com	98
unz.com	96
kp.ru	94
voltairenet.org	91
from-ua.com *	77
strategic-culture.org *	74
tass.ru	71

ria.ru	65
ussia-insider.com *	64
rbth.com	64
life.ru	58
rbc.ru	58
ussian.ru	57
southfront.org *	57
tvzvezda.ru	56
site.ru *	56
izvestia.ru	53
thesaker.is *	51
ussia.ru	50
ntv.ru	48
ussian.rt.com	47
vedomosti.ru	47

5 Highlighted Norwegian domains

Of the 37 anonymous Norwegian websites, only a few of them share links to the Russian-affiliated domains in our database at the time of this study. Domains may have had multiple owners and hosted different websites over the course of their lifetime. Majestic only captures data about the domain itself and does not track changes in domain ownership or the websites they host. Consequently, domains may have had a 'previous life' hosting websites that do not reflect its current use. These domains appear in our dataset because they were indexed by Majestic at some point and their ownership or content has since changed. Therefore, we do not provide the full list of the 37 anonymous domains in this report.

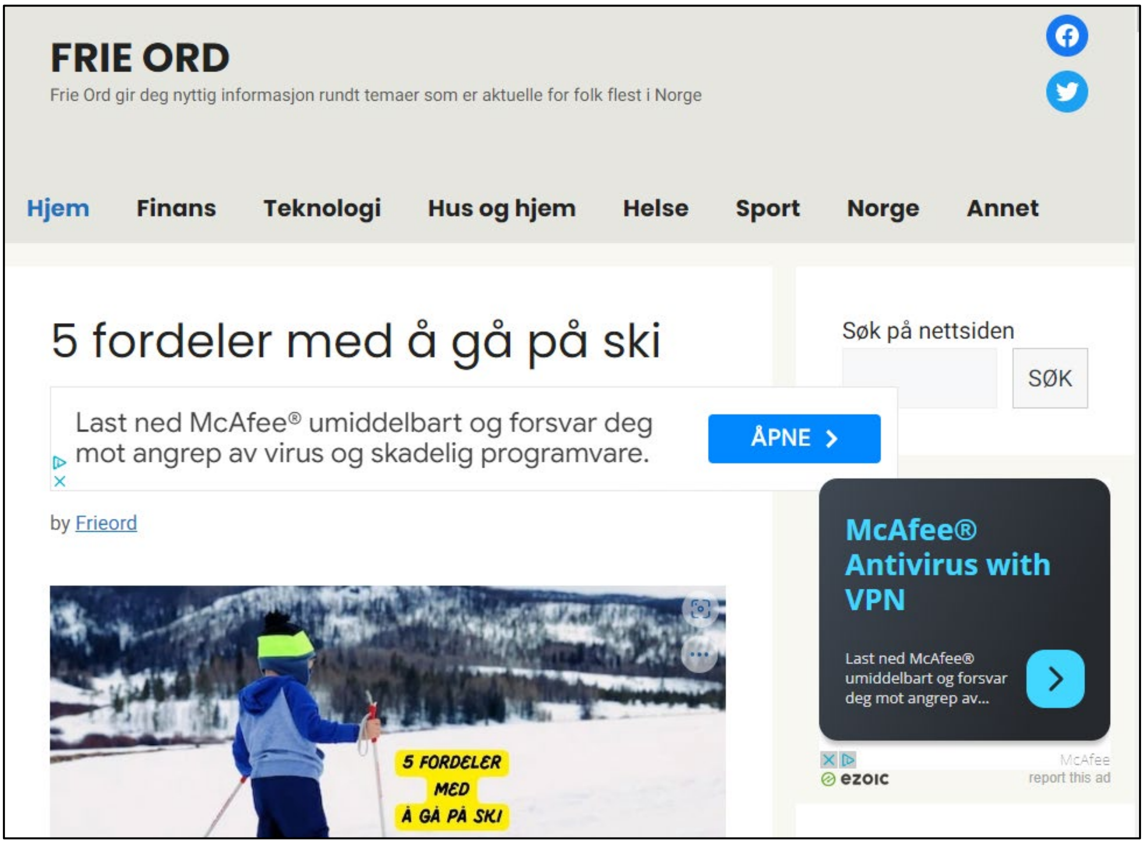
Instead, we have selected and further analysed four anonymous, Norwegian domains that stand out because they appear to be legitimate online news websites that follow editorial and journalistic principles, but display inauthentic content and feature fake reporters. Two of them are also part of a global network of inauthentic websites mimicking authentic news sites.

Whereas our initial assessment of the 693 filtered Norwegian domains was purposefully conducted only based on common digital literacy principles (i.e. looking for identifiable information about ownership and authors and in some cases performing simple Google searches to verify the information), we subjected the following, highlighted domains to further analysis. The purpose is to shed light on manipulative techniques we believe should be known to increase digital competency. *We neither claim nor imply that the following websites are connected to Russia.*

5.1 www.frieord.no

The domain *frieord.no* (“free words”) presents itself as a public news website with the stated purpose of providing “useful information on topics that most people have questions about” and displays a substantial amount of programmatic advertising (Figure 5.1). The domain has shared many links (4341) to Russian-affiliated domains, but does not seem to do so today. However, it stands out because it provides no information that can identify its owner and appears to feature authors with fake identities using common Norwegian names and artificially generated profile photos (GAN images). It provides a contact form and an email address (info@frieord.no). Upon contacting the website, the contact form yielded the response “The form was unable to submit. Please contact the administrator”. We received no reply to our email.

The domain has a specific Google Analytics (Universal Analytics) ID: UA-240676160. We have identified at least 25 other domains which use this same Google Analytics ID. Many of these appear to use GAN generated profile images. Researching these sites is beyond the scope of this report, which is why we do not list them.



FRIE ORD

Frie Ord gir deg nyttig informasjon rundt temaer som er aktuelle for folk flest i Norge

5 fordeler med å gå på ski

Last ned McAfee® umiddelbart og forsvar deg mot angrep av virus og skadelig programvare.

by Frieord

Last ned McAfee® umiddelbart og forsvar deg mot angrep av...

Figure 5.1 Screenshot of www.frieord.no, here with the top story “5 advantages to cross-country skiing”. Date of capture: 16 June, 2023.

Shares 4341 links to Russian sources

The domain *frieord.no* has shared a total of 4341 links to sources in the Russian disinformation and propaganda ecosystem, including rt.com (1859 links), zerohedge.com (1652 links), unz.com (599 links), sputniknews.com (210 links), globalresearch.ca (9 links), southfront.org (6 links), sott.net (2 links), strategic-culture.org (2 links) and voltairenet.org (2 links). A manual scan suggests that the website does not seem to link to Russian sources today, and that the identified link sharing may date back to a website that was previously hosted on the same domain. Data collected by the Internet Archive's Time Machine[27], shows that the domain hosted an "alternative news" site from 2014 and that the website currently hosted on the domain appeared in 2022. However, we include the current website here because it appears to have fake author profiles using artificially generated portraits.

Appears to use AI generated photos of authors

While the domain does not seem to share links to Russian sources today, it appears to feature inauthentic properties that makes it relevant to describe in this report from a digital literacy standpoint.

The website publishes content by three allegedly Norwegian journalists with common Norwegian names. By closer inspection, their portraits appear to be GAN generated (Generative Adversarial Network), meaning that they bear hallmarks of portraits created using artificial intelligence.[28] While this evaluation is subject to minor uncertainties, applying several detection methods[29] reveal that the pictures contain several features that are common for GAN generated faces. The profile pictures of three authors on frieord.no are shown in Figure 5.2.

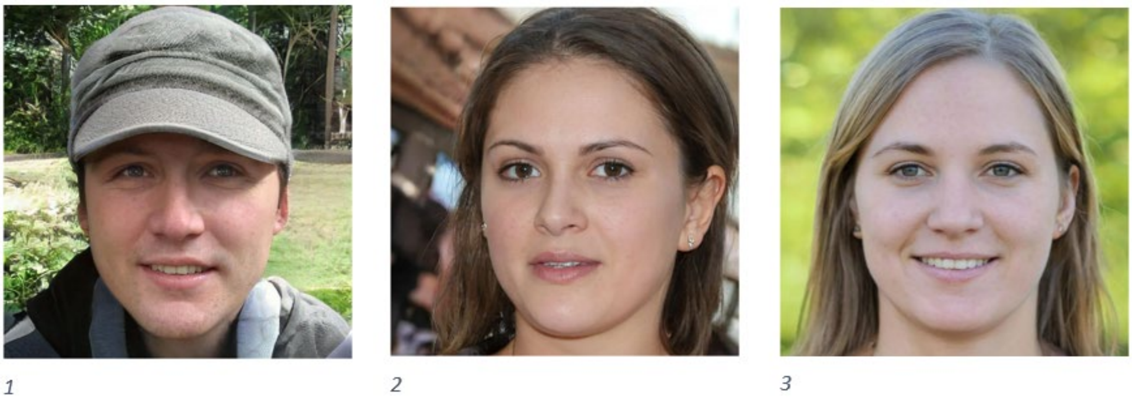


Figure 5.2 The portraits of the three authors on frieord.no contain features typical of artificially generated images.

[27] https://web.archive.org/web/20141116050831/https://www.frieord.no/

[28] Tanvi Arora & Rituraj Soni. (2021). *A review of techniques to detect the GAN-generated fake images*. In: Generative Adversarial Networks for Image-to-Image Translation (2021), pp. 125-159. Academic Press. Source: https://www.sciencedirect.com/science/article/abs/pii/B978012823519500004X

[29] Wang, X., Guo, H., Hu, S., Chang, M. & Lyu, S. (2023) *GAN-generated Faces Detection: A Survey and New Perspectives.* 26th European Conference on Artificial Intelligence (ECAI 2023). Source: https://arxiv.org/abs/2202.07145

All pictures contain features that represent physical inconsistencies, including:

- Misshaped or partly missing ears
- Unnatural structure in clothing
- Oddly stretched background
- Exaggerated borders and inconsistencies between depicted faces and backgrounds
- Pixelated or unnatural skin texture

The picture below (Figure 5.3) contains some of the clearest indicators of GAN faces; the texture of the hat and clothes appears unnatural, the border between the person and the background seems somewhat glitched in places, one ear appears to be partly missing, and the trees and branches in the background look unnaturally flat and tangled.

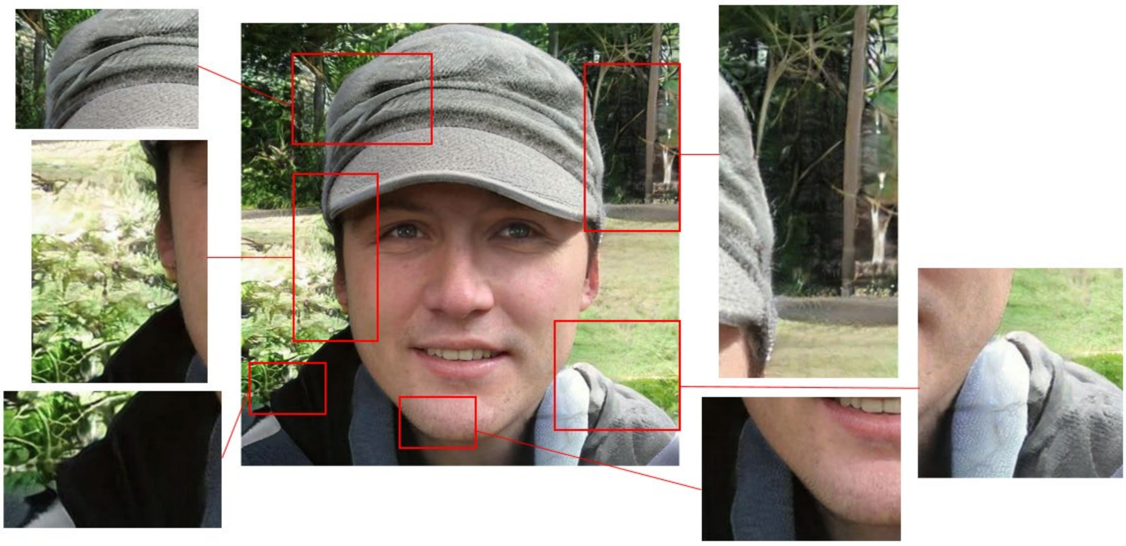


Figure 5.3 GAN indicators on the profile picture of author 1.

The other pictures contain similar indicators. In the picture of author 2 (Figure 5.4), the lines in the background do not align, and what appears to be a house (to the left of the face) is weirdly stretched and crooked in relation to the face. Furthermore, the earring on the right ear bleeds into the right cheek, the upper part of the other ear has a dark line across it and some hair seems to disappear unnaturally.

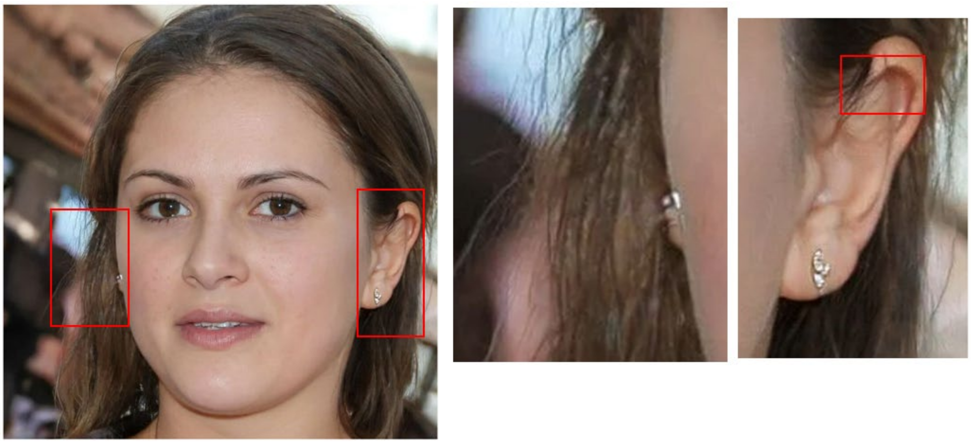


Figure 5.4 GAN indicators on the profile picture of author 2.

In the profile photo of author 3 the indicators are more subtle (Figure 5.5). There appears to be a few glitches in the hair and in the transition from the hair to the background, and the right ear seems misshaped. The strongest indicator, however, is the central alignment of the eyes.

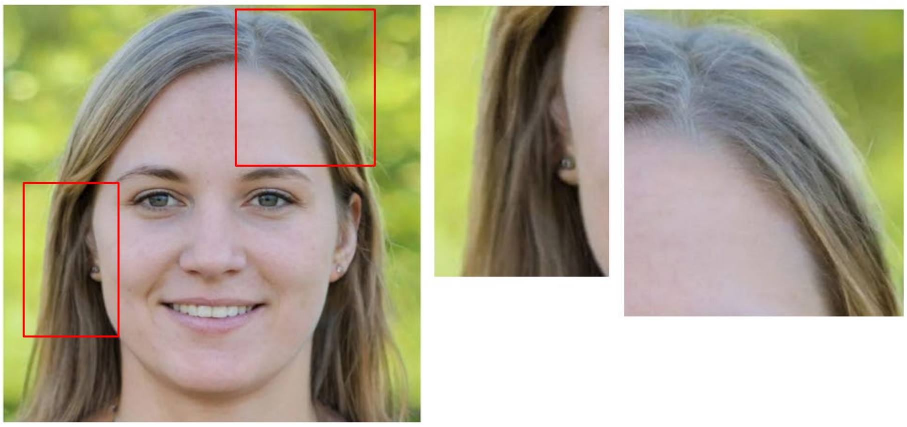


Figure 5.5 GAN indicators on the profile picture of author 3.

Glitches around the ears, eyes, and teeth are regarded as tell-tale indicators of GAN generated faces[30]. Additionally, blurred backgrounds and central eye alignment are common in such images. When increasing the transparency and layering the pictures on top of each other, it becomes clear that the eyes on all three images are centrally aligned, suggesting that the images are highly likely artificially generated (Figure 5.6).

[30] Stanford Internet Observatory (2022). *Unheard Voice: Evaluating five years of pro-Western covert influence operations*. Source: https://fsi.stanford.edu/publication/unheard-voice-evaluating-five-years-pro-western-covert-influence-operations-takedown

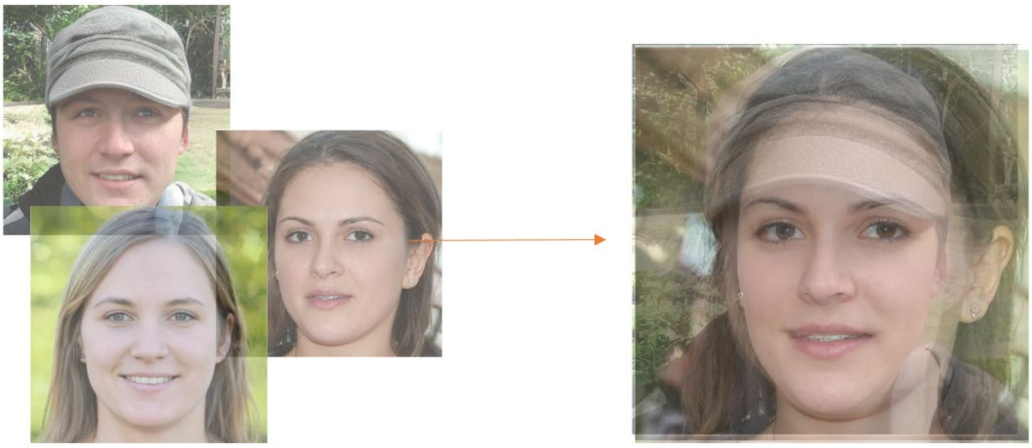


Figure 5.6 Increasing the transparency of all three images and layering them on top of each other shows that the eyes are centrally aligned.

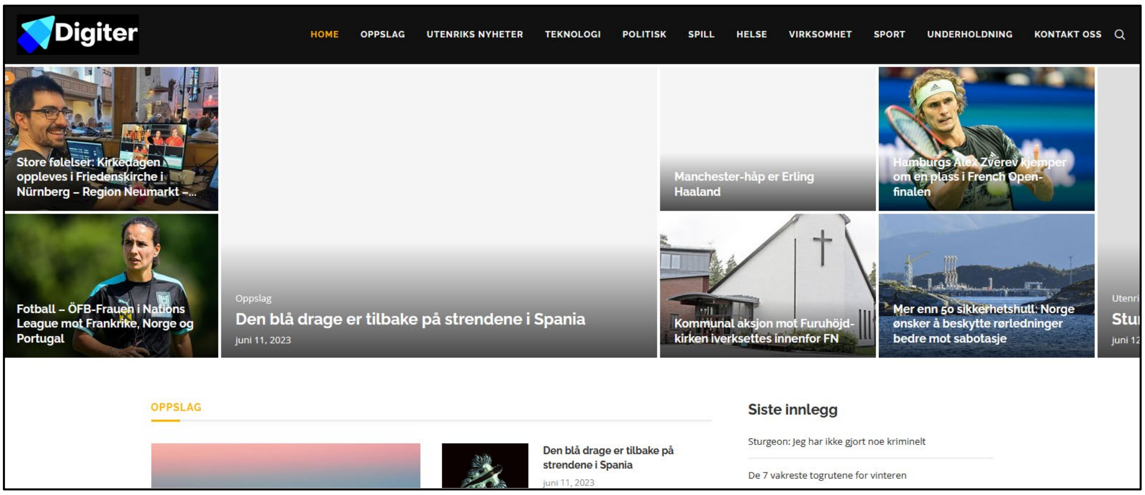
Concluding remarks

The website frieord.no has no identifiable owner, contains a significant amount of programmatic advertising and displays inauthentic properties in terms of what is highly likely fake authors with artificially generated profile photos. Its Google Analytics ID connects it to at least 25 other websites that run programmatic advertising, many of which display similar use of GAN generated photos. We have not further analysed the website nor the network to which it is connected, as this is beyond the scope of our study. The purpose of identifying it here is to describe what appears to be an inauthentic, anonymous Norwegian website that is part of a network designed to exploit programmatic advertising through posing as legitimate, informational websites. We argue that, from a digital literacy perspective, Internet users should be able to exercise sound judgment on websites' credibility, and that the use of fake authors with GAN generated profile photos is illegitimate and misleading.

5.2 www.digiter.no

The website *digiter.no* appears to be an online Norwegian newspaper covering topics such as politics, technology, sports, and entertainment and contains programmatic advertising (Figure 5.77). The domain has shared only 4 links to Russian sources; snanews.de (3 links) and tvzvezda.ru (1 link).

It stands out because it auto-copies and auto-translates articles from other sources and presents them as its own, its editorial staff have false identities and it provides no information about the website's owner beyond the email address Mileskarl05@gmail.com. The email address connects *digiter.no* to a global network of inauthentic and anonymously administered news sites in multiple languages that is further described in the following - as well as in the next case (5.3 kontrast1.no). We did not receive a reply to our email when contacting the website.



Store følelser: Kirkedagen oppleves i Friedenskirche i Nürnberg – Region Neumarkt –...

Manchester-håp er Erling Haaland

Hamburgs Alex Zverev kjemper om en plass i French Open-finalen

Fotball – ÖFB-Frauen i Nations League mot Frankrike, Norge og Portugal

Oppslag
Den blå drage er tilbake på strendene i Spania
juni 11, 2023

Kommunal aksjon mot Furuhöjd-kirken iverksettes innenfor FN

Mer enn 50 sikkerhetshull: Norge ønsker å beskytte rørledninger bedre mot sabotasje

Utenri
Stu
juni 12

OPPSLAG

Den blå drage er tilbake på strendene i Spania
juni 11, 2023

Siste innlegg

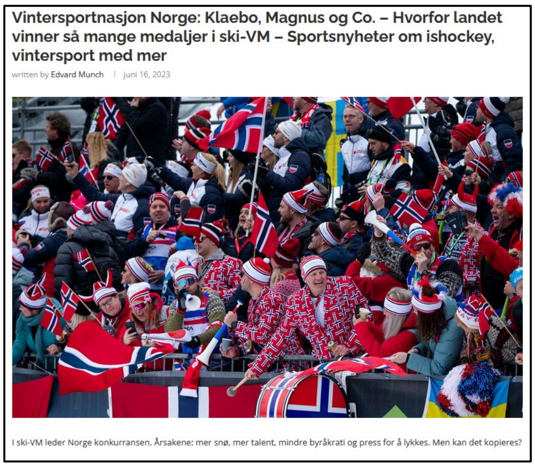
Sturgeon: Jeg har ikke gjort noe kriminelt

De 7 vakreste togrutene for vinteren

Figure 5.7 Screenshot of digiter.no. Date of capture: 16 June 2023.

Auto-copied and auto-translated content

Articles on *digiter.no* reveal many errors and strange wordings that indicate that the text has been auto-translated to Norwegian from other languages. For example, one article – curiously written by a *digiter.no* author named "Edvard Munch" (Norwegian painter) – shows many signs of auto-translation illustrative of the language in the articles we sampled. A simple Google search reveals that the article is copied and translated from a German news site, *Allgaüer Zeitung* (Figure 5.8and Figure 5.9).

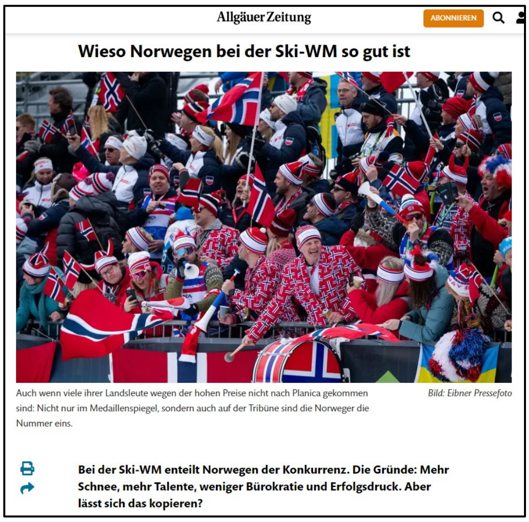


Vintersportnasjon Norge: Klaebo, Magnus og Co. – Hvorfor landet vinner så mange medaljer i ski-VM – Sportsnyheter om ishockey, vintersport med mer

written by Edvard Munch | juni 16, 2023

I ski-VM leder Norge konkurransen. Årsakene: mer snø, mer talent, mindre byråkrati og press for å lykkes. Men kan det kopieres?

Figure 5.8 Article on digiter.no. Source: https://digiter.no/vintersportnasjon-norge-klaebo-magnus-og-co-hvorfor-landet-vinner-sa-mange-medaljer-i-ski-vm-sportsnyheter-om-ishockey-vintersport-med-mer-2/



Wieso Norwegen bei der Ski-WM so gut ist

Auch wenn viele ihrer Landsleute wegen der hohen Preise nicht nach Planica gekommen sind: Nicht nur im Medaillenspiegel, sondern auch auf der Tribüne sind die Norweger die Nummer eins.

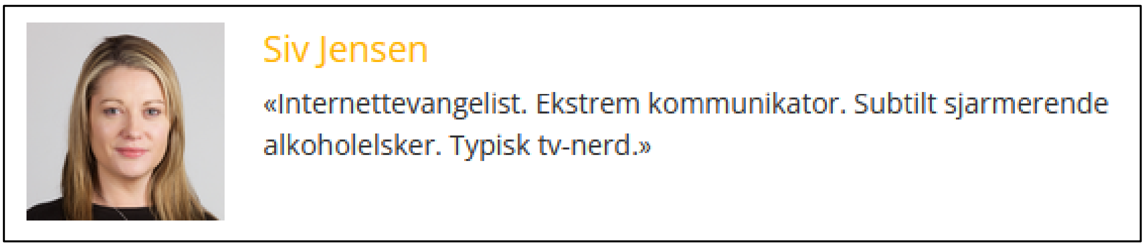
Bild: Eibner Pressefoto

Bei der Ski-WM enteilt Norwegen der Konkurrenz. Die Gründe: Mehr Schnee, mehr Talente, weniger Bürokratie und Erfolgsdruck. Aber lässt sich das kopieren?

Figure 5.9 Original article on Allgaüer Zeitung. Source: https://www.allgaeuer-zeitung.de/sport/die-norweger-machens-unkomplizierter_arid-542961

Fake identities on editorial staff

Articles on *digiter.no* feature bylines with the name and a photo of the author, but no contact details. Two authors are named "Liv Ullmann" (famous Norwegian actress) and "Edvard Munch" (famous Norwegian painter, deceased). Examples of other names are "Siv Jensen" (famous Norwegian former politician, but also a common name) and "Ashley Olsen" (famous American former actress, i.e. "the Olsen twins"). Their profile photos are taken from other sources. For example, here is the bio of the journalist "Siv Jensen" (Figure 5.10).

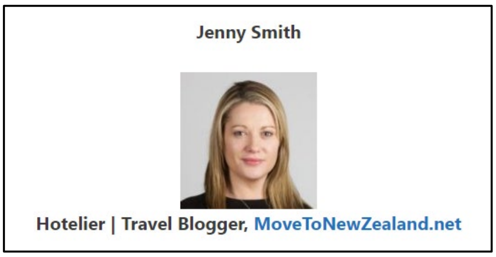


Siv Jensen

«Internettevangelist. Ekstrem kommunikator. Subtilt sjarmerende alkoholelsker. Typisk tv-nerd.»

Figure 5.10 Screenshot of the bio of the journalist "Siv Jensen". English translation: "Internet evangelist. Extreme communicator. Subtly charming alcohol lover. Typical TV nerd."

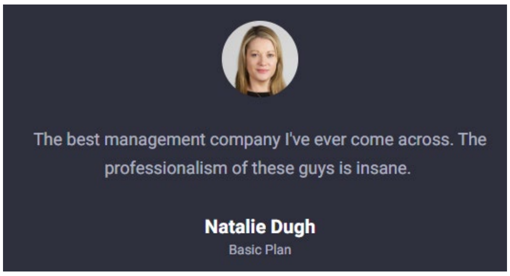
Using Google's reverse image search, we find the same profile photo used on at least 11 other websites in various countries and languages, and on at least three different LinkedIn profiles. In all instances, the photo is displayed with different names (or no name), none of which are "Siv Jensen". See two examples below in Figure 5.11 and Figure 5.12.



Jenny Smith

Hotelier | Travel Blogger, MoveToNewZealand.net

Figure 5.11 Screenshot from the domain "upjourney.com", using the same photo with the name "Jenny Smith". Source:https://upjourney.com/whatis-the-difference-between-a-hotel-motel-inn

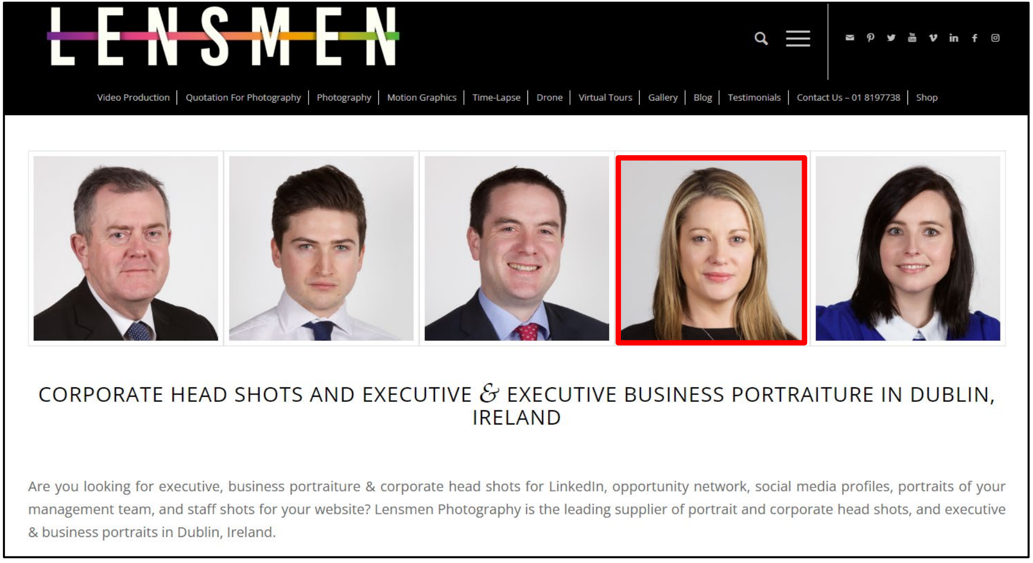


The best management company I've ever come across. The professionalism of these guys is insane.

Natalie Dugh
Basic Plan

Figure 5.12 Screenshot from the domain "slow2ventures.co", using the same photo with the name "Natalie Dugh". Source: https://slow2ventures.co/

Google's reverse image search places the likely origin of the photo on an Irish video and photography production company where it is showcased as one of several examples of their business portraits (Figure 5.13).

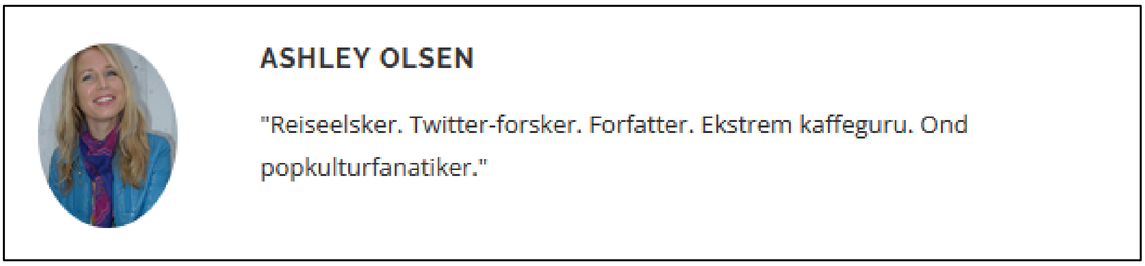


CORPORATE HEAD SHOTS AND EXECUTIVE & EXECUTIVE BUSINESS PORTRAITURE IN DUBLIN, IRELAND

Are you looking for executive, business portraiture & corporate head shots for LinkedIn, opportunity network, social media profiles, portraits of your management team, and staff shots for your website? Lensmen Photography is the leading supplier of portrait and corporate head shots, and executive & business portraits in Dublin, Ireland.

Figure 5.13 The profile photo of «Siv Jensen" as it appears on the Irish video and photography company "Lensmen". Source: https://www.lensmen.ie/editorial-photography/business-portraiture/

Another journalist on *digiter.no* is “Ashley Olsen”, presented with this bio (Figure 5.14):

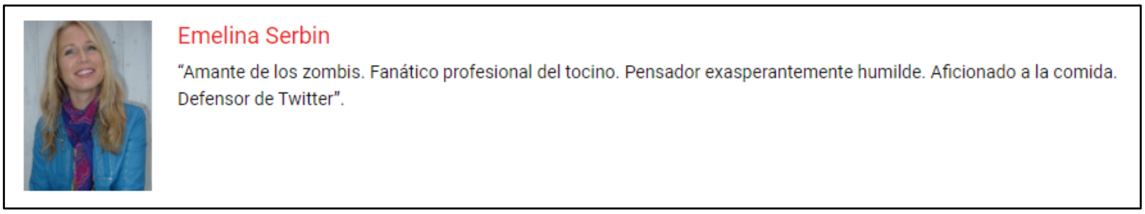


ASHLEY OLSEN

"Reiseelsker. Twitter-forsker. Forfatter. Ekstrem kaffeguru. Ond popkulturfanatiker."

Figure 5.14 Screenshot of the bio of the journalist “Ashley Olsen”. English translation (ours): “Travel lover. Twitter researcher. Author. Extreme coffee guru. Evil pop culture fanatic.”

Using Google’s reverse image search, we found the same photo on the bio of a journalist on the Chilean domain *publinoticias.cl*, under the name “Emelina Serbin” (Figure 5.15). The website on the domain has a similar design to *digiter.no* and lists Mileskarl05@gmail.com as the owner, just as with *digiter.no*.

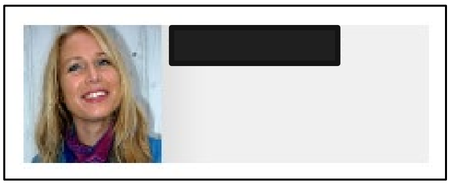


Emelina Serbin

“Amante de los zombis. Fanático profesional del tocino. Pensador exasperantemente humilde. Aficionado a la comida. Defensor de Twitter”.

Figure 5.15 On the domain publicnoticias.cl the same photo is used to present its journalist “Emilina Serbin”. English translation of the bio: “Zombie lover. Professional bacon fanatic. Maddeningly humble thinker. Foodie. Twitter Defender”.

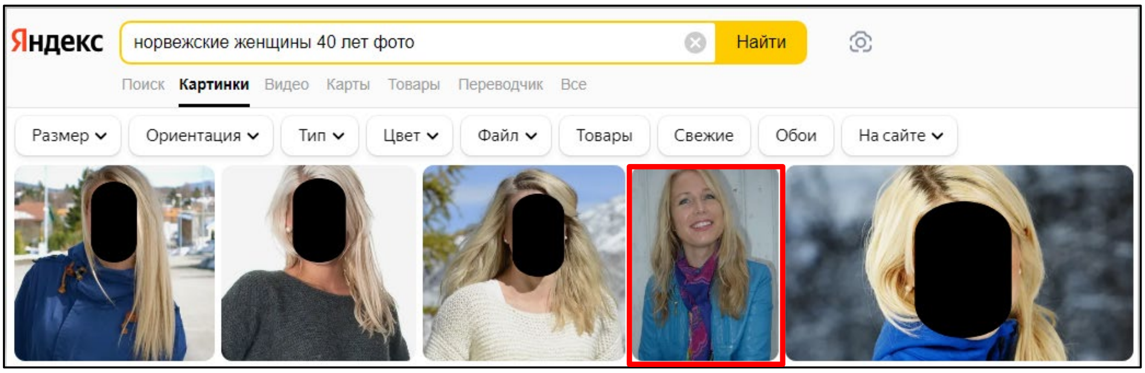
We found a cropped version of the same profile photo used on the website of an established Norwegian cultural event, where the person is identified by her real name (which is not Ashley Olsen or any of the other) and occupation (Figure 5.16). We omit further information from this report for privacy reasons.



[redacted]

Figure 5.16 Screenshot from the website of the cultural event where the person is identified by her real name and occupation.

Google's reverse image search attributes the origin of the photo to the Russian Yandex' image search, in a category called "Norwegian women 40 years old photo" (Google translated from Russian "норвежские женщины 40 лет фото") (Figure 5.17). From here, the photo links to its source, an online forum on the domain *theapricity.com*, in a category called "Norwegian women". We were not able to identify how it made its way to the forum.



норвежские женщины 40 лет фото

Figure 5.17 Screenshot of the Russian Yandex' image search engine displaying the full-size version of the profile photo used by digiter.no

Part of an international network

A Google search for "Mileskarl05@gmail.com" shows the same email address as the owner of more than two hundred other websites in multiple languages, including Norwegian, English, Spanish, French, Japanese, Greek, Czech, Italian, Portuguese and Polish – to mention a few.

The websites display similar design and content categories, giving the impression that they are legitimate news sites. In addition, many of them share a curious trait – a mismatch between the domain name and the content. For example, two of the other 13 Norwegian websites we identified connected with the email address Mileskarl05@gmail.com, are *vossblues.no* and *betzykrisenter.no*. The names do not indicate that they are news sites, but rather websites for a blues club[31] and a centre for people in need[32] – which they in fact were when they were established. A plausible explanation is that the entity behind Mileskarl05@gmail.com directly, and/or via proxies, purchases expired domains with pre-existing backlinks as these domains will perform better in web searches.[33]

As we took a closer look at the next website, described in the following case (5.3), we discovered that *digiter.no* and the other websites of the "Mileskarl05@gmail.com" network are part of an even larger network. The most probable reason for the mismatch between the domain names and the content on their websites, is that acquiring legacy domains is a search engine optimization tactic which increases online visibility and reach.

[31] Former/original website: https://web.archive.org/web/20041130095218/http://www.vossblues.no/

[32] Former/original website: https://web.archive.org/web/20110921014252/http://www.betzykrisenter.no/

[33] We have not looked into domain ownership, as that is beyond the scope of this report.

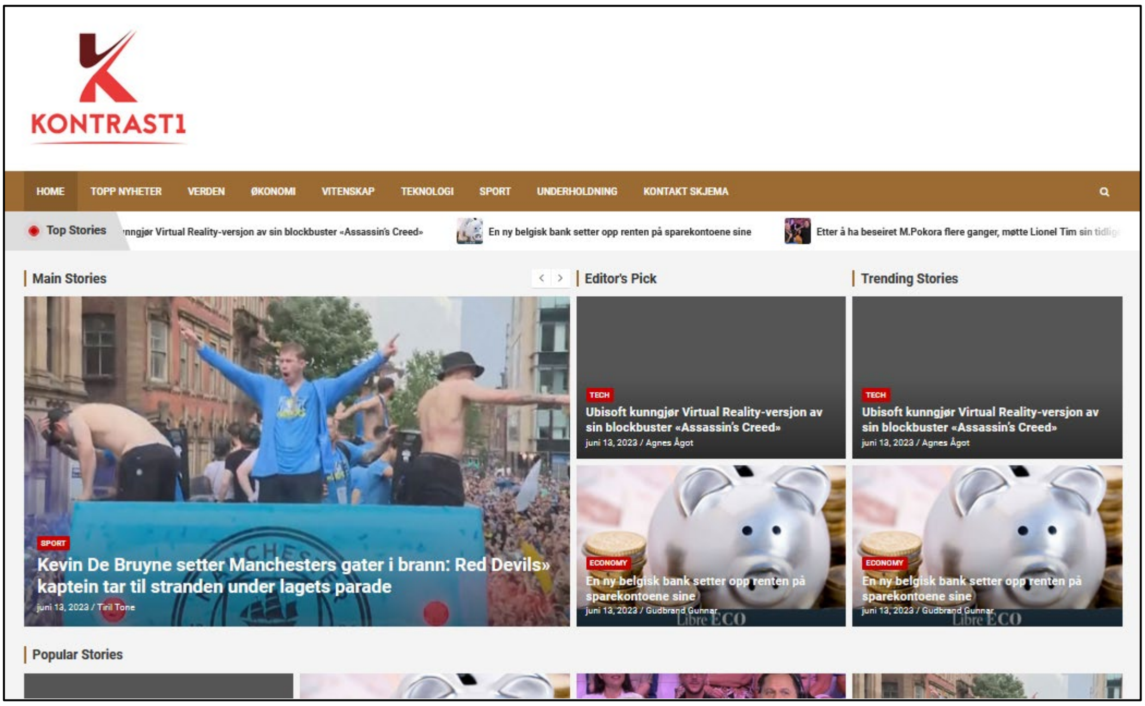
Concluding remarks

We choose to describe *digiter.no* in this report because it may shed light on a phenomenon that we believe people should know about from a digital literacy standpoint. The domains linked to Mileskarl05@gmail.com seem to be inauthentic websites mimicking legitimate news sites in local languages. They operate with fake authors and auto-copy and auto-translate their articles from other sources. It is likely that the entity behind Mileskarl05@gmail.com directly and via proxies purchases expired domains (such as *vossblues.no*) with pre-existing backlinks, meaning it will perform better on web searches – which explains the mismatch between the domain names and the content. This is supported by findings connected to the website we describe in the following subchapter, 5.3 kontrast1.no.

While the network provides ample opportunity for generating advertising revenue and the potential to spread any information to a larger audience through mimicking authentic news sites in local languages, further analysis of the network is beyond the scope of this report. However, as the following case will describe, this network is part of an even larger network.

5.3 www.kontrast1.no

The domain *kontrast1.no* ("the contrast") appears to be an online Norwegian newspaper (Figure 5.18). It shares only three links Russian sources, all from tass.com. It stands out because it provides no information about ownership or authors beyond the email address powerhayden58@gmail.com, which links it to a global network of websites similar to that of Mileskarl05@gmail.com described in the previous case example of *digiter.no*. However, the website does not contain programmatic advertising. Its authors have fake identities using Norwegian names (some generic and some only with oddly constructed double first names), and no other information that can verify their identity. The format of the authors' bios also resembles the author bios in the Mileskarl05@gmail.com network. Based on our correspondence with powerhayden58@gmail.com, described later in this case, we conclude that they are indeed part of the same network made up of 443 inauthentic websites in 32 languages.



nngjør Virtual Reality-versjon av sin blockbuster «Assassin's Creed»

En ny belgisk bank setter opp renten på sparekontoene sine

Etter å ha beseiret M.Pokora flere ganger, møtte Lionel Tim sin tidlig

Kevin De Bruyne setter Manchesters gater i brann: Red Devils» kaptein tar til stranden under lagets parade
juni 13, 2023 / Tiril Tone

Ubisoft kunngjør Virtual Reality-versjon av sin blockbuster «Assassin's Creed»
juni 13, 2023 / Agnes Ågot

Ubisoft kunngjør Virtual Reality-versjon av sin blockbuster «Assassin's Creed»
juni 13, 2023 / Agnes Ågot

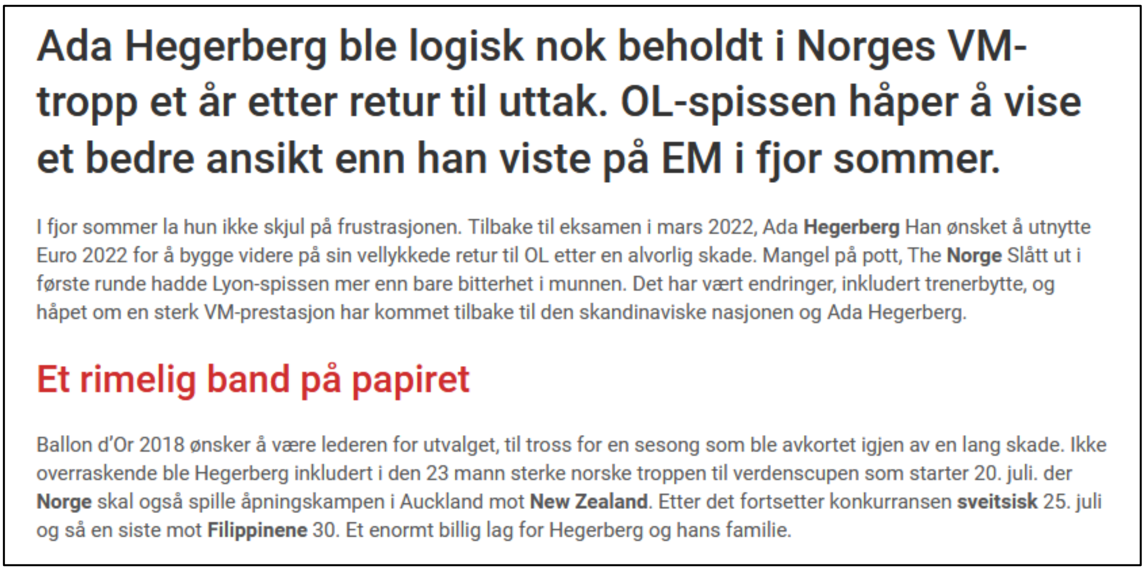
En ny belgisk bank setter opp renten på sparekontoene sine
juni 13, 2023 / Gudbrand Gunnar

En ny belgisk bank setter opp renten på sparekontoene sine
juni 13, 2023 / Gudbrand Gunnar

Figure 5.18 Screenshot of www.kontrast1.no. Date of capture: 16 June 2023.

Auto-copied and auto-translated content

As with *digiter.no*, the content of *kontrast1.no* seems to be auto-translated to Norwegian. For example, the first sentence of the "about us" page reads: "Kontrast1 Lyser over de siste og viktige gjennombruddene i de siste topphistoriene, politikk, teknologi, oppstart, helse og vitenskap via lettleste innlegg". The sentence appears to have been auto-translated from another language, possibly English, based on the incorrect use of the word "Lyser" (with a capital L and possibly translated from "sheds light") and the word "oppstart" (possibly translated from "startups”). A random sampling of articles shows that they bear clear hallmarks of automated translation. Figure 5.19 provides one of several examples in which the article is full of incomplete and incoherent sentences and incorrect use of words and phrases typical for auto translation.



Ada Hegerberg ble logisk nok beholdt i Norges VM-tropp et år etter retur til uttak. OL-spissen håper å vise et bedre ansikt enn han viste på EM i fjor sommer.

I fjor sommer la hun ikke skjul på frustrasjonen. Tilbake til eksamen i mars 2022, Ada **Hegerberg** Han ønsket å utnytte Euro 2022 for å bygge videre på sin vellykkede retur til OL etter en alvorlig skade. Mangel på pott, The **Norge** Slått ut i første runde hadde Lyon-spissen mer enn bare bitterhet i munnen. Det har vært endringer, inkludert trenerbytte, og håpet om en sterk VM-prestasjon har kommet tilbake til den skandinaviske nasjonen og Ada Hegerberg.

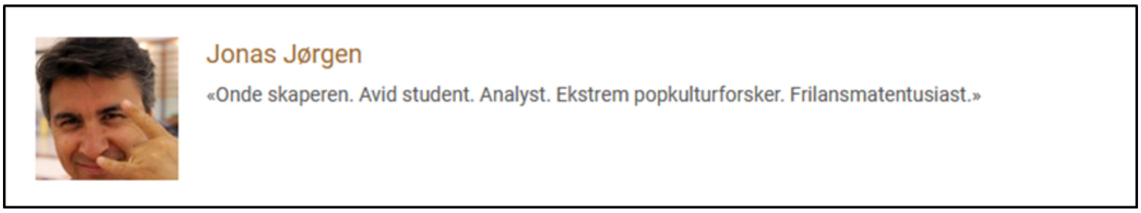
Et rimelig band på papiret

Ballon d’Or 2018 ønsker å være lederen for utvalget, til tross for en sesong som ble avkortet igjen av en lang skade. Ikke overraskende ble Hegerberg inkludert i den 23 mann sterke norske troppen til verdenscupen som starter 20. juli. der **Norge** skal også spille åpningskampen i Auckland mot **New Zealand**. Etter det fortsetter konkurransen **sveitsisk** 25. juli og så en siste mot **Filippinene** 30. Et enormt billig lag for Hegerberg og hans familie.

Figure 5.19 Example of content from kontrast1.no that bears hallmarks of auto-translation.

Fake identities on editorial staff

Articles on *kontrast1.no* include bylines with the name and a photo of the author, but no contact details. The names are either common Norwegian names that are difficult to verify without extensive research, or an unusual combination of two Norwegian first names, such as “Jonas Jørgen”, “Gudbrand Gunnar”, and “Tiril Tone”. Their bios are odd and grammatically and/or linguistically incorrect and resemble those of *digiter.no* and the websites of the Mileskarl05@gmail.com network previously described. The journalists’ profile pictures are taken from other sources. For example, here is the bio of the journalist “Jonas Jørgen” (Figure 5.20).

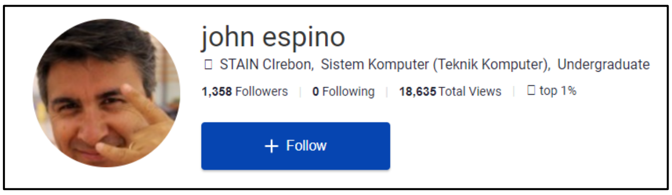


Jonas Jørgen

«Onde skaperen. Avid student. Analyst. Ekstrem popkulturforsker. Frilansmatentusiast.»

Figure 5.20 The bio of the journalist «Jonas Jørgen” is linguistically incorrect and thus difficult to translate into equally incorrect English. It reads “Evil creator. Avid student. Analyst. Extreme pop culture researcher. Freelance food enthusiast”.

Using Google reverse image search, we find the same profile photo used on at least 30 other websites, two LinkedIn profiles and four Twitter/X profiles, all with different names. Here are two examples from other websites (Figure 5.21 and Figure 5.22).



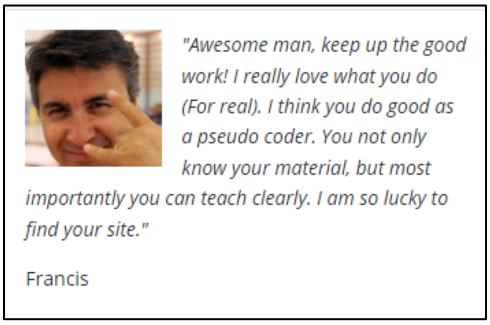
john espino

□ STAIN Clrebon, Sistem Komputer (Teknik Komputer), Undergraduate

1,358 Followers | 0 Following | 18,635 Total Views | □ top 1%

+ Follow

Figure 5.21 The same profile photo used for the profile of "John Espino" on the domain www.academia.edu. Source: https://staincirebon.academia.edu/johnespino



"Awesome man, keep up the good work! I really love what you do (For real). I think you do good as a pseudo coder. You not only know your material, but most importantly you can teach clearly. I am so lucky to find your site."

Francis

Figure 5.22 And here he is as "Francis" providing a testimonial on the domain https://dev.ilovecoding.org/ .

International network

A simple google search for "powerhayden58@gmail.com" shows the same email address as the owner of at least 37 other, similar websites in several languages, including Norwegian, German, French, Arabic, Italian, Spanish, Portuguese, Dutch and Polish – to mention some. As such, the network bears resemblance to that of Mileskarl05@gmail.com described in the previous case study of *digiter.no*. While this network also provides ample opportunity to spread any kind of information to local populations by mimicking authentic news sites, further research into the actors and the purpose behind the network is beyond the scope of this report. We did, however, contact the website to ask about our findings and received information that enabled us to identify many more websites as well as the connection between the "powerhayden58" and the "Mileskarl05" networks.

Contacting the website confirms global network

Upon contacting the website in Norwegian using the email address powerhayden58@gmail.com, we received what appears to be a standardized, English reply with the signature of a Power Hayden, CEO of Power Press Agency. No other identifiable information was provided. The email did not answer our query about inauthentic properties. Instead, it was a sales pitch stating that they offer "sponsored press release and guest posting opportunities" as well as "sponsored link insertion opportunities" on their "editorial and non-editorial websites", of which most are "Google News-approved old publisher properties, providing instant indexing of the article". According to the email, this "helps to increase your search engine ranking, as news websites have high trust with Google and signals will be passed on to you".

The email offered a link to a Google doc that lists a total of 443 websites in 32 languages in the "powerhayden58" network. The websites are broken down into three categories, of which 327 sites are described as "editorial sites" in 22 languages (including *kontrast1.no*), 110 sites are described as "guest posting sites" in 15 languages and six sites are described as "non-editorial websites/niche blogs" in English. While the "editorial sites" do not seem to display programmatic advertising, the "guest posting sites" and "non-editorial websites/niche blogs" do.[34]

The website *kontrast1.no* is the only Norwegian website in the "editorial sites" category. However, in the "guest posting sites" category, we find additional 13 Norwegian websites[35], including *digiter.no*, which we described in the previous case example (5.2), and the other Norwegian sites we identified as part of the Mileskarl05 network. While the visual similarities between the websites of the powerhayden58 and Mileskarl05 networks suggested a connection between the two, the list confirms that they are the same network.

We never received a reply to our questions about the use of fake author profiles and auto-translated text. Because the network appears to be inauthentic using illegitimate methods such as fake journalists and auto-copying/translation of content from established news sources, we performed Google searches in an attempt to confirm the authenticity of "Power Hayden" and the entity "Power Press Agency" before describing our findings in this report. We were not able to confirm either.

Based on the names, the two Gmail addresses and the domain names it is possible to find out more about the network and the actors involved than what is described in this report. However, our focus is on the identified websites from a digital literacy point-of-view and not on identifying the actors behind it.

[34] The full list of 443 domains may be requested for legitimate research or journalistic purposes.

[35] The 13 Norwegian websites are: Yttersiden.no, Topshineauto.no, Thecoolgirl.no, Norskmatkultur.no, Securmarksykkel.no, Kjaerra.no, Betzykrisesenter.no, Citra2010oslo.no, Vossblues.no, Nyematoghelse.no, Norskoffroadteknikk.no, Easydisplay.no, Digiter.no.

Concluding remarks

Based on our findings and the information provided in the email from powerhayden58@gmail.com, it is reasonable to conclude that the 443 websites that list Mileskarl05@gmail.com and powerhayden58@gmail.com as the only contact information comprise a global network of inauthentic, anonymous sites in 32 languages.[36] They mimic authentic news media and blogs to look credible, and the content is either auto-copied and auto-translated from other news sources or paid for by third parties and presented as editorial articles that appear to have been written by real journalists, but who have fake identities.

Based on the findings and the information provided by powerhayden58@gmail.com, it is highly likely that the purpose of the network is to make money from displaying paid-for, third-party content disguised as legitimate news articles, and to create advertising revenue from drawing Internet users to their websites. A plausible explanation for the mismatch between domain names and the content on the websites they host, is that the entity behind the network directly and/or via proxies purchases expired domains with pre-existing backlinks, meaning it will perform better on web searches.

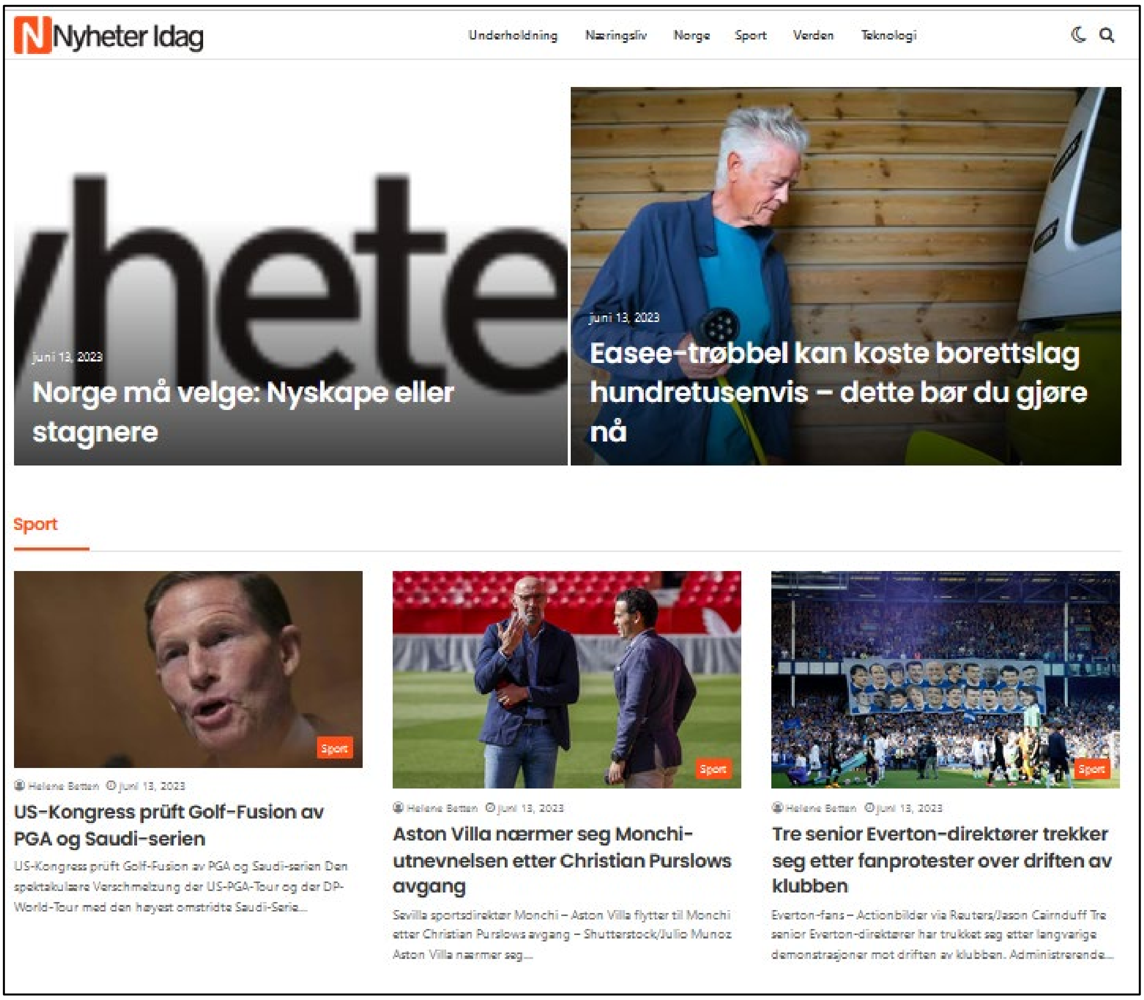
The network provides ample opportunity to spread any kind of (dis)information disguised as news from editorial news outlets to large audiences in local languages globally. While the primary purpose of the network seems to be financial gain, it offers an established, global infrastructure that gives anyone, including malign actors, an opportunity to plant and spread paid-for content with the aim of exerting covert influence directed at target groups in multiple languages.

[36] Languages: Arabic, Brazilian Portuguese, Bulgarian, Chinese, Chinese (Hong Kong), Czech, Dutch, English, French, German, Greek, Hindi, Hungarian, Indonesian, Italian, Japanese, Korean, Latvian, Lithuanian, Norwegian, Polish, Portuguese, Romanian, Russian, Serbian, Slovakian, Spanish, Swedish, Thai, Turkish, Ukrainian, Vietnamese

5.4 www.nyheteridag.news

The website *nyheteridag.news* ("news today") has shared only 3 links to the official Russian source tass.com, and provides no information about the website's owner. The domain stands out because it appears to be a proper Norwegian online newspaper while it hides its owner's identity (Figure 5.23). It provides the email address info@nyheteridag.news as the only contact information. When contacting the website, we did not get a response.

It seems to auto-translate and auto-copy news from other news sources, change the wording slightly and present the news using a byline with only a generic Norwegian name and no other information or contact details that can verify the author's identity. The website contains programmatic advertising.



juni 13, 2023

Norge må velge: Nyskape eller stagnere

juni 13, 2023

Easee-trøbbel kan koste borettslag hundretusenvis – dette bør du gjøre nå

Sport

Helene Betten juni 13, 2023

US-Kongress prüft Golf-Fusion av PGA og Saudi-serien

US-Kongress prüft Golf-Fusion av PGA og Saudi-serien Den spektakulære Verschmelzung der US-PGA-Tour og der DP-World-Tour med den høyest omstridte Saudi-Serie...

Helene Betten juni 13, 2023

Aston Villa nærmer seg Monchi-utnevnelsen etter Christian Purslows avgang

Sevilla sportsdirektør Monchi – Aston Villa flytter til Monchi etter Christian Purslows avgang – Shutterstock/Julio Munoz Aston Villa nærmer seg...

Helene Betten juni 13, 2023

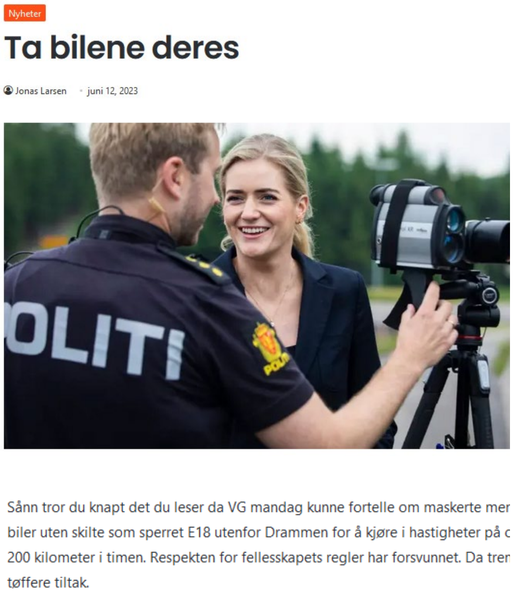
Tre senior Everton-direktører trekker seg etter fanprotester over driften av klubben

Everton-fans – Actionbilder via Reuters/Jason Cairnduff Tre senior Everton-direktører har trukket seg etter langvarige demonstrasjoner mot driften av klubben. Administrerende...

Figure 5.23 Screenshot of www.nyheteridag.news. Date of capture: 16 June 2023.

Auto-copied and auto-translated content

As with other websites described in this report, *nyheteridag.news* also appears to copy articles from other sources and present them as its own, including changing the byline to that of its own editorial staff – whose identity cannot be verified. A random sampling of articles shows that articles are copied from authentic, Norwegian news sites, such as *Dagsavisen.no*, *tv2.no*, and *NRK.no*. However, the website does not copy the articles in verbatim, but changes the wording slightly. See illustrative example below, where *nyheteridag.news* has taken an opinion piece from *Dagsavisen.no*, kept the same headline and photograph, but performed slight alterations to the text. Here, we have selected a random paragraph to illustrate this (Figure 5.24 and 5.25).

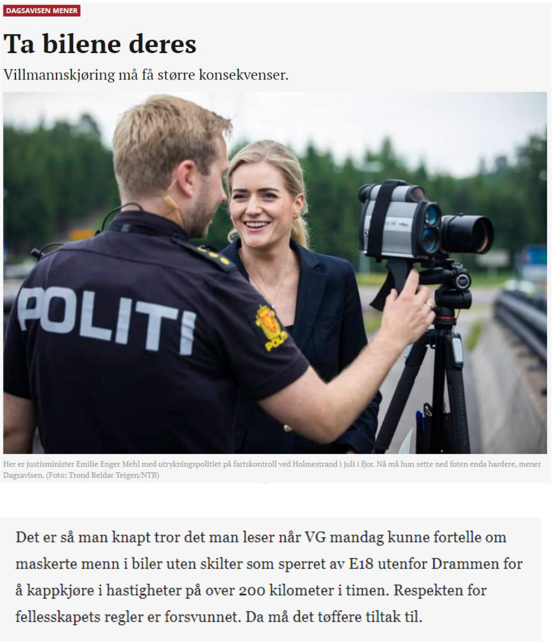


Nyheter

Ta bilene deres

Jonas Larsen • juni 12, 2023

Sånn tror du knapt det du leser da VG mandag kunne fortelle om maskerte mer
biler uten skilte som sperret E18 utenfor Drammen for å kjøre i hastigheter på c
200 kilometer i timen. Respekten for fellesskapets regler har forsvunnet. Da trer
tøffere tiltak.



DAGSAVISEN MENER

Ta bilene deres

Villmannskjøring må få større konsekvenser.

Her er justisminister Emilie Enger Mehl med utrykningspolitiet på fartskontroll ved Holmestrand i juli i fjor. Nå må hun sette ned foten enda hardere, mener Dagsavisen. (Foto: Trond Reidar Teigen/NTB)

Det er så man knapt tror det man leser når VG mandag kunne fortelle om maskerte menn i biler uten skilter som sperret av E18 utenfor Drammen for å kappkjøre i hastigheter på over 200 kilometer i timen. Respekten for fellesskapets regler er forsvunnet. Da må det tøffere tiltak til.

Figure 5.24

Screenshot of headline and 1st paragraph of news article on nyheteridag.info, 12 June 2023, copied from Dagsavisen's opinion piece the same day (displayed on the right). The wording in the story is slightly changed from the original and presented under a byline with a non-verifiable generic name, "Jonas Larsen". Source: https://nyheteridag.news/ta-bilene-deres/

Figure 5.25

Screenshot of headline and 1st paragraph of the original opinion piece on Dagsavisen.no, 12 June 2023. Source: https://www.dagsavisen.no/debatt/leder/2023/06/12/ta-bilene-deres/

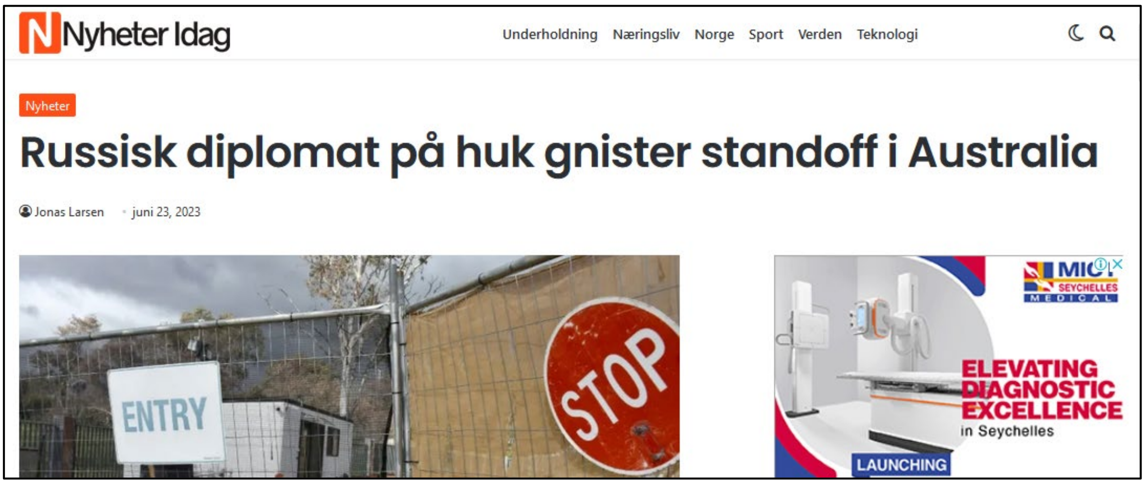
The alterations from the original text found on *Dagsavisen.no* is obvious, as illustrated in Table 5.1.

Table 5.1 Examples of slight text alterations between original and copied article

Original Norwegian text on Dagsavisen.no	Copied and altered text on nyheteridag.news
Det er så man knapt tror det man leser (…)	Sånn tror du knapt det du leser (…)
(…) for å kappkjøre i hastigheter (…)	(…) for å kjøre i hastigheter (…
Respekten for fellesskapets regler er forsvunnet.	Respekten for fellesskapets regler har forsvunnet.
Da må det tøffere tiltak til.	Da trengs tøffere tiltak.

While this is inarguably plagiarism, *nyheteridag.news* does provide the name of the website from which it copied the content from (but not a link to the article itself). For example, at the bottom of the article referenced above, it says "Kilde: Dagsavisen" (Eng. "Source: Dagsavisen").

Nyheteridag.news displays articles from both Norwegian and international sources. In the latter case, the wording strongly indicates auto-translation. For example, in the article displayed below (Figure 5.26), the headline is nonsensical in Norwegian, and in the text itself the phrase "Dette er en del av den russiske lekeboken" strongly suggests auto-translation; "Lekeboken" makes no sense in Norwegian, and is highly likely auto-translated from English "playbook".



Russisk diplomat på huk gnister standoff i Australia

Jonas Larsen · juni 23, 2023

Figure 5.26 Article on nyheteridag.news where the headline is nonsensical in Norwegian, indicating that it has been auto-translated. Date of capture: 16 June 2023.

The provided source of this article is simply "via Bing". A quick Google search identifies the original article as an AFP article in English syndicated by numerous news sites.

Concluding remarks

The website on the domain *nyheteridag.news* displays inauthentic properties that include auto-copying and auto-translation of articles taken from other sources. These articles are then presented as original content by journalists whose identity cannot be verified. The wording in the auto-copied articles are slightly altered before published on *nyheteridag.news*, which makes it a little harder to find the original based on normal Google searches. Its programmatic advertising may suggest that its purpose is financial gain through displaying advertising, however this has not been analysed further. We did not identify a Google Analytics ID that could connect *nyheteridag.news* to other websites.

5.5 Summary and key takeaways

The four highlighted websites were chosen not because they have shared links to Russian-affiliated sources, but because they display inauthentic properties that we believe should be made publicly known from a digital literacy perspective. Upon investigating them further, two of them led us to uncover a global network of 443 inauthentic websites in 32 languages that mimic legitimate news sites. Of these, 14 are in Norwegian.

Creating inauthentic news sites in order to generate revenue from programmatic advertising is not a new phenomenon, but may become a bigger challenge with the availability of generative AI. NewsGuard has to date (30 November 2023) identified 566 AI-generated news and information sites operating with little to no human oversight.[37] As stated by NewsGuard, "in many cases, the revenue model for these websites is programmatic advertising under which the ad-tech industry delivers ads without regard to the nature or quality of the website. As a result, top brands are unintentionally supporting these sites. Unless brands take steps to exclude untrustworthy sites, their ads will continue to appear on these types of sites, creating an economic incentive for their creation at scale." (ibid.).

While we have no indications that the global network we uncovered has any other purpose than financial gain, it does represent an established infrastructure that may be exploited by malign actors to influence target groups in multiple languages. Further research into the content and reach of this network is beyond the scope of our study, but is recommended.

[37] NewsGuard. (2023). *Tracking AI-enabled Misinformation: 566 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools.* Source: https://www.newsguardtech.com/special-reports/ai-tracking-center/

6 Limitations

6.1.1 Limitations of a link-based approach

This study's link-based approach has an obvious blind spot. We are only able to uncover domains that share links to specific domains. We have looked for backlinks to 2686 Russian-affiliated domains, but the total number of these type of domains is unknown and constantly changing as domains are taken down and new ones appear. With our approach, we are not able to identify anonymous Norwegian domains that link to Russian-affiliated domains which we are not aware of.

One possible course of action to mitigate this limitation could be to gather all the *outbound* links of the domains we have uncovered with Majestic. This could give us a new list of potentially unknown and relevant domains to investigate in-links to. This, however, was outside of the scope of this study.

Another intrinsic limitation of the link-based approach is that we only follow *links* to uncover shared content. Online content is obviously shared in numerous other ways, such as copy-paste of text, translation of text into other languages, and sharing of images, sound and video etc.

6.1.2 Limitations of using Majestic SEO tool

We have utilized a commercial SEO tool (Majestic) which was built and designed for other purposes than ours. As such, there are bound to be some limitations with its use.

Tracking domain changes over time
Majestic does not track changing ownership of a domain over time. This means that a domain can have had a "previous life" that differs from the current version of a domain. Put simply, a domain that has existed for some time may have hosted a string of different websites with different owners and content over the course of its lifetime. As far as we have been able to identify, there is no way to differentiate between different "versions" of domains over time using Majestic (this can be done manually using the Internet Archive Wayback Machine[38]). The fact that websites continuously change over time and that links can both appear and disappear is a known challenge with these types of studies[39].

Identifying time
Using Majestic alone, we are not able to identify when a web page was created, or when a hyperlink was published. This would require continuously crawling and indexing to be able to timestamp changes over time. Majestic operates with a timestamp for when a web page was

[38] Internet Archive Wayback Machine: http://web.archive.org/

[39] Enrique Orduña-Malea & Rodrigo Costas. (2021). *Link-based approach to study scientifc software usage: the case of VOSviewer.* Scientometrics. Source: https://link.springer.com/article/10.1007/s11192-021-04082-y

crawled and indexed, which provides a proxy to understanding when content might have been changed.

Cut-offs for returned referring domains

We have used the Majestic API command *GetRefDomains*[40], which has some limitations. When requesting referring domains from Majestic, a maximum of 100 000 results are returned per input domain. We set the parameter "OrderBy1" = 11, which means that the domains are ordered by the number of links to the source in our database.

This means that for domains with more than 100 000 referring domains, "only" the 100 000 domains with the most links to the source domain in our database will be returned. 28 of the 2686 Russian-affiliated domains returned 100 000 results. We do not know what the actual number is. However, for 20 of the 28 domains that returned 100 000 referring domains, the referring domain with the lowest number of matched links was 3 or more. This means that the cut-off affects linking domains which had 1 or 2 backlinks to one of the 20 domains. We therefore evaluate the loss of data to be insignificant for our overall findings.

Most of the sources in our database are not affected by this cut-off (2658 of 2686). Domains affected by the cut-off limitation for one Russian-affiliated source in our database might still be identified by its links to *other* sources.

Lack of context

We only get aggregated numbers for how many times a domain has backlinked to our initial list of domains. As we do not get an overview of the specific URLs which contain links to our list of Russian-affiliated sources, we have no way of knowing the context in which these links have been published. The context might be that the website that shares links to Russian-affiliated domains agrees or disagrees with the content they are linking to, or for example that it is being sarcastic about it. Also, links can have been published in the comments section of a website (if that exists), meaning that the owners or authors of the domain did not intend to link to the content. This has, however, not been a limitation for the analysis of identified Norwegian sites in this study, as we have manually assessed the filtered results.

Distinguishing between subdomains

We have not been able to distinguish between subdomains of popular blogging platforms, like blogspot.com, livejournal.com and wordpress.com. This seems to be a limitation with the Majestic API command which we have used. This means that we may be underreporting on the various sites that could be backlinking to our list of Russian-affiliated domains.

Website structures

The structure of a website which is crawled and indexed might also pose some challenges for a SEO tool provider like Majestic. Websites are designed and coded in myriad ways and known issues like asynchronous loading of content might prevent the crawler from finding all the links.

[40] https://developer-support.majestic.com/api/commands/

Also, backlinks might be indexed multiple times, as they figure in headers, menus and footers which are used in multiple webpages on one single domain.

6.1.3 Other limitations

Because we base our qualitative analysis on common digital literacy practices, i.e. evaluating the information that is visible on the website itself, we may have overlooked websites that have been meticulously crafted by an advanced actor to look legitimate.

The qualitative analysis revealed a gray zone in which at least five websites could not be fully regarded as anonymous, nor could they be regarded as fully verified. For example, some sites display limited or ambiguous information about the actor behind them, but features authentic and identifiable Norwegians as contributing authors. It is possible that Russia or Russian-affiliated actors can establish such websites to amplify authentic Norwegian voices that support Russian narratives. However, when in doubt we have chosen not to identify such sites to avoid the risk of labelling the exercise of free speech problematic.

7 Conclusion

The purpose of this exploratory study was to gain insight into if and how content from Russia's official state media websites (such as RT, Sputnik etc.) and proxy websites is spread to domestic audiences in countries outside Russia, more specifically Norway. Our approach has been based on three specific research questions. In the following we provide summarized answers to them, based on our findings, before closing off with concluding remarks.

RQ 1: Do Norwegian domains share links to Russian official and proxy domains in the Russian disinformation and propaganda ecosystem?

Yes. We identified **1356** Norwegian domains that do so, when defining "Norwegian domains" as .no domains and any domain that has more than 50% Norwegian content (bokmål and nynorsk) with a language detection confidence level of more than 70%. **976** of them have shared 3 or more links. When filtering out the domains with websites that do not seem to be online today, the number is **693**.

RQ 2: Do any of them do so anonymously?

Yes. Based on qualitative analysis, we identified **37** Norwegian websites that share links to Russian-affiliated domains while not providing verifiable information about the website's owner or authors. However, most of them do not seem to share links to Russian-affiliated websites today. They appear in our dataset because of previous websites hosted on the same domain. We have therefore not included a list of the domains in this report.

RQ 3: Do any of them contain inauthentic or manipulated content?

Yes. In this report we have highlighted four websites that display various inauthentic and manipulative traits, including fake author profiles with stolen portraits, GAN generated profile images, auto-copying and auto-translation of content and artificially created photos of reporters. Two of them mimic authentic news outlets and are connected to a global network of 443 similar inauthentic and anonymous websites in 32 languages, of which 14 websites are Norwegian. (See chapter 5 – Highlighted Norwegian Domains). We neither suggest nor imply that any these websites have a connection to Russia. The two Norwegian websites that appeared in our dataset, and are connected to the global network, have shared a very limited number of links to Russian-affiliated domains. We have not checked link-sharing from the other 441 domains in the network they are a part of.

Concluding remarks

There may be many reasons why someone would create anonymous Norwegian websites and use them to share links to the Russian disinformation and propaganda ecosystem. There may also be many reasons why these websites contain inauthentic or manipulated content. Neither suggest a connection to Russia in and of itself. However, it *may*. Local language distribution of

Russian narratives through proxy sources and the use of fake identities is an integral part of Russia’s strategy for information influence.

In the qualitative part of our research, we approached the websites just as any Internet user would, based on common digital literacy practice, i.e. looking at the content on the website itself. We have considered a website to be anonymous when it does not provide verifiable or truthful information about its owner or authors.

While we did find 37 anonymous and five partly verifiable Norwegian domains that have shared links to Russian-affiliated domains, we have not found evidence to suggest that they have - or have had - a connection to Russia, nor do we imply that they do. Answering such questions would require deeper research beyond the scope and framework of this study.

The four anonymous Norwegian websites highlighted for deeper analysis in this report all appear in our dataset because they have shared links to Russian-affiliated domains. However, the reason they are highlighted is because they employ inauthentic methods that we choose to expose from a transparency and digital literacy standpoint. Upon investigating them further, two of them led us to uncover a global network of 443 inauthentic, anonymous websites in 32 languages that mimic legitimate news sites. 14 of the websites are Norwegian. We argue that this is problematic from a digital literacy standpoint. We also argue that this, and similar, networks represent a potential infrastructure for (dis)information distribution that may be exploited by malign actors for the purpose of exerting covert influence and manipulations.

Based on our findings, it is reasonable to conclude that the Russian propaganda and disinformation ecosystem does not seem to have a foothold in Norway *through anonymous, Norwegian proxy sites that share hyperlinks to other Russian-affiliated domains* during the time of this study. However, content from Russian-affiliated websites may be shared by anonymous Norwegian websites in other ways than through link-sharing, for example by translating articles without linking to the source. Such sites will not be identified by this study.

References

Dudek et al. (2021). *Co-link analysis as a monitoring tool: A webometric use case to map the web relationships of research projects. In: Proceedings of the 18th International Conference on Scientometrics & Informetrics (2021)*, pp. 339-344. Source: https://arxiv.org/abs/2110.04251

Flem, S. S. & Molnes, G. (2020). *Slik spres russisk propaganda i norske alternative medier.* Faktisk.no. Source: https://www.faktisk.no/artikler/06epg/slik-spres-russisk-propaganda-i-norske-alternative-medier

Giles, Keir et al. (2018). *Russian Reflexive Control.* Royal Military College of Canada. Source: https://publications.gc.ca/site/fra/9.881883/publication.html

Global Engagement Center. (2020). *GEC Special Report: Pillars of Russia's disinformation and propaganda ecosystem. U.S. State Department.* Source: https://www.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia%E2%80%99s-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf (p. 3).

Hanley, Hans W. A. et al. (2022). *No Calm in the Storm: Investigating QAnon Website Relationships.* Stanford University. Source: https://ojs.aaai.org/index.php/ICWSM/article/view/19293/19065

Internet Archive Wayback Machine: http://web.archive.org/

Jalal, S., Sutradhar, B., Sahu, K., Mukhopadhyay, P., & Biswas, S. (2015). *Search Engines and Alternative Data Sources in Webometric Research: An Exploratory Study. In: DESIDOC Journal of Library & Information Technology, 35(6)*. Source: https://doi.org/10.14429/djlit.35.6.8883

Majestic SEO tool: https://majestic.com/

Medietilsynet. (2021). *Stop, think, check: How to expose fake news and misinformation. Medietilsynet.* Source: https://www.medietilsynet.no/digitale-medier/kritisk-medieforstaelse/stop-think-check-en/

Nato. MC 0628 *NATO Military policy on Strategic Communications*

NewsGuard. (2023). *Tracking AI-enabled Misinformation: 566 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools.* Source: https://www.newsguardtech.com/special-reports/ai-tracking-center/

Nimmo, Ben (2023). *Detailed report: Taking down coordinated inauthentic behavior from Russia and China.* Meta. Source: https://about.fb.com/wp-content/uploads/2022/11/CIB-Report_-China-Russia-Sept-2022.pdf

OpenFacto (2023). *InfoRos' historical networks of influence.* Source: https://openfacto.fr/2023/01/16/inforos-historical-networks-of-influence/

Orduña-Malea, E. & Costas, R. (2021). *Link-based approach to study scientifc software usage: the case of VOSviewer. Scientometrics.* Source: https://link.springer.com/article/10.1007/s11192-021-04082-y

Pak, Chankyung et al. (2020). *Intermedia Reliance and Sustainability of Emergent Media: A Large-Scale Analysis of American News Outlets' External Linking Behaviors. International Journal of Communications.* Source: https://ijoc.org/index.php/ijoc/article/view/13040

Rodríguez, Belén Carrasco. (2020). *Information Laundering in the Nordic-Baltic region.* Nato Strategic Communications Center of Excellence. Source: https://stratcomcoe.org/publications/information-laundering-in-the-nordic-baltic-region/26

Sehgal, Vibhor et al. (2021). *Mutual Hyperlinking Among Misinformation Peddlers.* University of California, Berkeley. Source: https://arxiv.org/abs/2104.11694

Shenoy, A., Prabhu, A. (2016). *SEO Hub: Utilities and Toolsets. In: Introducing SEO.* Apress, Berkeley, CA. Source: https://doi.org/10.1007/978-1-4842-1854-9_10

Sivertsen et al. (2022). *Uønsket utenlandsk påvirkning? – kartlegging og analyse av stortingsvalget 2021.* Forsvarets forskningsinstitutt. Source: https://www.ffi.no/publikasjoner/arkiv/uonsket-utenlandsk-pavirkning-kartlegging-og-analyse-av-stortingsvalget-2021

Stanford Internet Observatory (2022). *Unheard Voice: Evaluating five years of pro-Western covert influence operations.* Source: https://fsi.stanford.edu/publication/unheard-voice evaluating-five-years-pro-western-covert-influence-operations-takedown

Suad Kunosić, Denis Čeke and Enver Zerem. (2018). *Advantages and Disadvantages of the Webometrics Ranking System. In: Scientometrics Recent Advances.* IntechOpen. Source: https://www.intechopen.com/chapters/67912

Tanvi Arora & Rituraj Soni. (2021). *A review of techniques to detect the GAN-generated fake images. In: Generative Adversarial Networks for Image-to-Image Translation (2021)*, pp. 125-159. Academic Press. Source: https://www.sciencedirect.com/science/article/abs/pii/B978012823519500004X

Varghese, M. & Lawrance, Reejo M. (2019). *Webometric Studies A Review of Literature. In: ILIS Journal of Librarianship and Informatics Vol. 2, No. 1.*, pp. 91 – 101. Source: https://www.academia.edu/43766555/Webometric_Studies_A_Review_of_Literature_Reejo_M_Lawrance

Waissbluth, Elliott et al. (2021). *Domain-Level Detection and Disruption of Disinformation.* University of California, Berkeley. Source: https://arxiv.org/abs/2205.03338

Wang, X., Guo, H., Hu, S., Chang, M. & Lyu, S. (2023). *GAN-generated Faces Detection: A Survey and New Perspectives. 26th European Conference on Artificial Intelligence (ECAI 2023)*. Source: https://arxiv.org/abs/2202.07145

About FFI

The Norwegian Defence Research Establishment (FFI) was founded 11th of April 1946. It is organised as an administrative agency subordinate to the Ministry of Defence.

FFI's mission

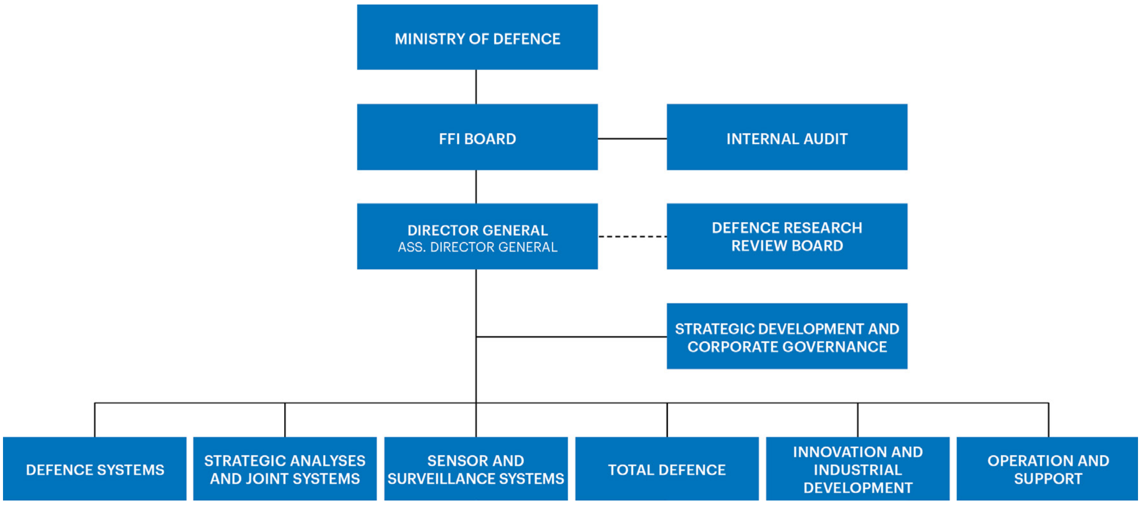
FFI is the prime institution responsible for defence related research in Norway. Its principal mission is to carry out research and development to meet the requirements of the Armed Forces. FFI has the role of chief adviser to the political and military leadership. In particular, the institute shall focus on aspects of the development in science and technology that can influence our security policy or defence planning.

FFI's vision

FFI turns knowledge and ideas into an efficient defence.

FFI's characteristics

Creative, daring, broad-minded and responsible.


MINISTRY OF DEFENCE
FFI BOARD
INTERNAL AUDIT
DIRECTOR GENERAL
ASS. DIRECTOR GENERAL
DEFENCE RESEARCH REVIEW BOARD
STRATEGIC DEVELOPMENT AND CORPORATE GOVERNANCE
DEFENCE SYSTEMS
STRATEGIC ANALYSES AND JOINT SYSTEMS
SENSOR AND SURVEILLANCE SYSTEMS
TOTAL DEFENCE
INNOVATION AND INDUSTRIAL DEVELOPMENT
OPERATION AND SUPPORT

Forsvarets forskningsinstitutt (FFI)
Postboks 25
2027 Kjeller

Besøksadresse:
Kjeller: Instituttveien 20, Kjeller
Horten: Nedre vei 16, Karljohansvern, Horten

Telefon: 91 50 30 03
E-post: post@ffi.no
ffi.no

Norwegian Defence Research Establishment (FFI)
PO box 25
NO-2027 Kjeller
NORWAY

Visitor address:
Kjeller: Instituttveien 20, Kjeller
Horten: Nedre vei 16, Karljohansvern, Horten

Telephone: +47 91 50 30 03
E-mail: post@ffi.no
ffi.no/en